

Soft-Pion Theorems for Large Scale Structure

Bart Horn, Lam Hui and Xiao Xiao

*Institute for Strings, Cosmology and Astroparticle Physics (ISCAP) and Physics Department
Columbia University, New York, NY 10027, USA*

Abstract

Consistency relations – which relate an N-point function to a squeezed (N+1)-point function – are useful in large scale structure (LSS) because of their non-perturbative nature: they hold even if the N-point function is deep in the nonlinear regime, and even if they involve astrophysically messy galaxy observables. The non-perturbative nature of the consistency relations is guaranteed by the fact that they are symmetry statements, in which the velocity plays the role of the soft pion. In this paper, we address two issues: (1) how to derive the relations systematically using the residual coordinate freedom in the Newtonian gauge, and relate them to known results in ζ -gauge (often used in studies of inflation); (2) under what conditions the consistency relations are violated. In the non-relativistic limit, our derivation reproduces the Newtonian consistency relation discovered by Kehagias & Riotto and Peloso & Pietroni. More generally, there is an infinite set of consistency relations, as is known in ζ -gauge. There is a one-to-one correspondence between symmetries in the two gauges; in particular, the Newtonian consistency relation follows from the dilation and special conformal symmetries in ζ -gauge. We probe the robustness of the consistency relations by studying models of galaxy dynamics and biasing. We give a systematic list of conditions under which the consistency relations are violated; violations occur if the galaxy bias is non-local in an infrared divergent way. We emphasize the relevance of the adiabatic mode condition, as distinct from symmetry considerations. As a by-product of our investigation, we discuss a simple fluid Lagrangian for LSS.

1 Introduction

In studies of large scale structure (LSS), we are familiar with the consequence of *linearly realized* symmetries. For instance, consider a time-independent spatial translation:

$$\vec{x} \rightarrow \vec{x} + \vec{\epsilon}, \quad (1)$$

where $\vec{\epsilon}$ is constant in time and space. For a small $\vec{\epsilon}$, we can think of its effect on the density field δ as:

$$\delta \rightarrow \delta + \Delta\delta \quad \text{where} \quad \Delta\delta \equiv -\vec{\epsilon} \cdot \vec{\nabla} \delta. \quad (2)$$

Notice how the change in the field $\Delta\delta$ depends *linearly* on the field i.e. the symmetry is linearly realized. Its consequence for an N-point correlation function is well known:

$$\Delta\langle\delta(\vec{x}_1)\dots\delta(\vec{x}_N)\rangle = -\vec{\epsilon} \cdot \sum_{a=1}^N \vec{\nabla}_a \langle\delta(\vec{x}_1)\dots\delta(\vec{x}_N)\rangle = 0, \quad (3)$$

where $\vec{\nabla}_a$ stands for $\partial/\partial\vec{x}_a$. This is nothing other than an infinitesimal form of the statement that the correlation function is translationally invariant: $\langle\delta(\vec{x}_1)\dots\delta(\vec{x}_N)\rangle = \langle\delta(\vec{x}_1 - \vec{\epsilon})\dots\delta(\vec{x}_N - \vec{\epsilon})\rangle$. Assuming the initial conditions were translationally invariant, the correlation function at late times has the same property since the dynamics itself does.

What is somewhat less familiar in LSS is the consequence of *nonlinearly realized* symmetries. By this, we mean that the symmetry involves transforming the field of interest in a nonlinear way. In this paper, a field that transforms this way will turn out to be the velocity field \vec{v} , i.e.

$$\Delta\vec{v} = \Delta_{\text{lin.}}\vec{v} + \Delta_{\text{nl.}}\vec{v}, \quad (4)$$

where $\Delta_{\text{lin.}}\vec{v}$ depends linearly on the fluctuating variable \vec{v} , while $\Delta_{\text{nl.}}\vec{v}$ does not depend on the fluctuating field, or any fluctuating field for that matter – in this sense, it might be more precise to call $\Delta_{\text{nl.}}\vec{v}$ sublinear. An example is a *time-dependent* spatial translation:

$$\vec{x} \rightarrow \vec{x} + \vec{n}(\eta), \quad (5)$$

where \vec{n} depends on (conformal) time η but not space. It is evident that under such a transformation, in addition to the usual linear transformation $\Delta_{\text{lin.}}\vec{v} = -\vec{n} \cdot \vec{\nabla}\vec{v}$, the velocity experiences a nonlinear shift:

$$\Delta_{\text{nl.}}\vec{v} = \vec{n}', \quad (6)$$

where $'$ denotes a time derivative. It was pointed out recently by Kehagias & Riotto and Peloso & Pietroni [1, 2] (KRPP) that just such a time-dependent spatial translation is in fact a symmetry of the familiar system of equations for pressureless fluid coupled to gravity in the Newtonian limit: the continuity, Euler and Poisson equations. As pointed out by the same authors, the consequence of such a nonlinearly-realized symmetry is *not* the simple invariance of a general correlation function (such as in Eq. 3), but rather a relation between an (N+1)-point function and an N-point function:

$$\lim_{\vec{q} \rightarrow 0} \frac{\langle\vec{v}(\vec{q})\mathcal{O}_{\vec{k}_1}\mathcal{O}_{\vec{k}_2}\dots\mathcal{O}_{\vec{k}_N}\rangle^{c'}}{P_v(q)} \sim \langle\mathcal{O}_{\vec{k}_1}\mathcal{O}_{\vec{k}_2}\dots\mathcal{O}_{\vec{k}_N}\rangle^{c'}, \quad (7)$$

where P_v is the velocity power spectrum, the superscript c' denotes a *connected* correlation function with the overall delta function removed, and \mathcal{O} denotes some observable, which could be different for each \vec{k} .

This sort of *consistency relation*, which relates a squeezed (N+1)-point function to an N-point function, is well known in the context of single field inflation. The first example was pointed out by Maldacena [3] (see also [4]):

$$\lim_{\vec{q} \rightarrow 0} \frac{\langle\zeta_{\vec{q}}\zeta_{\vec{k}_1}\dots\zeta_{\vec{k}_N}\rangle^{c'}}{P_\zeta(q)} = - \left(3(N-1) + \sum_{a=1}^N \vec{k}_a \cdot \frac{\partial}{\partial\vec{k}_a} \right) \langle\zeta_{\vec{k}_1}\dots\zeta_{\vec{k}_N}\rangle^{c'}, \quad (8)$$

where ζ is the curvature perturbation. It arises from a spatial dilation symmetry, which is non-linearly realized on ζ . Recently, more non-linearly realized symmetries were uncovered, including the special conformal symmetry [5, 6] and in fact a whole infinite tower of symmetries [7] (H2K hereafter). Recent work has emphasized the non-perturbative nature of these consistency relations as Ward/Slavnov-Taylor identities [7–13]. They can be viewed as the cosmological analog of the classic soft-pion theorems, which relate a scattering amplitude with N particles to another one with the same set of particles plus a

soft Nambu-Goldstone boson. In the case of the strong interactions, the symmetry in question is chiral symmetry, which is spontaneously broken. The corresponding Nambu-Goldstone boson is the pion, which shifts nonlinearly under chiral symmetry. In the examples above, ζ and the velocity play the role of the pion – both shift nonlinearly under the respective symmetries.

The non-perturbative nature of the consistency relations, such as Eq. (7), makes them very interesting for LSS studies. The hard modes, with momenta labeled by \vec{k} , can be highly nonlinear and can even be astrophysically messy observables, such as galaxy density as opposed to mass density. The robustness of these consistency relations is not completely surprising: we are familiar with the statement that correlation functions should be translationally invariant (Eq. 3), even if they involve fluctuations that are highly nonlinear and/or astrophysically complex. The robustness stems from the fact that these statements, whether the analogs of Eq. (3) or (7), are derived from symmetries, linearly or nonlinearly realized. Nonetheless, the consistency relations from nonlinearly realized symmetries are less familiar in the context of LSS, and deserve further investigation.

First, we wish to probe the consistency relations by studying them in the context of concrete models of galaxy dynamics and biasing. An often instructive way to understand supposedly robust relations is to ask when they actually break down. We will delineate the assumptions underlying the consistency relations and study examples where they are violated. We will see that the consistency relations stand on three legs: the existence of nonlinearly realized symmetries, the single field initial condition, and the adiabatic mode condition. We emphasize especially the role of the last in the breakdown of consistency relations. Using the language of galaxy-biasing, we can phrase a general violation as the sign of a particular kind of non-local galaxy bias. As a by-product of our analysis, we also write down a Goldstone boson effective Lagrangian for a pressureless fluid coupled to gravity, similar to effective actions for quintessence [14, 15].

Second, we wish to establish the connection between the Newtonian consistency relation found by KRPP (Eq. 7), and the relativistic consistency relations in the context of inflation. How are they related? Given the infinite number of consistency relations known for inflation, could there be more such relations for the context of large scale structure? The inflationary consistency relations are usually phrased in unitary or ζ gauge, while the Newtonian gauge is more natural for LSS. We will show, not surprisingly, that there is a one-to-one correspondence between symmetries in the two gauges. The time-dependent spatial translation of KRPP (Eq. 5) turns out to correspond to the special conformal symmetry in ζ gauge. The full list of symmetries in the Newtonian gauge can be inferred from the known list in ζ gauge. The KRPP consistency relation is the only one with a non-trivial Newtonian/sub-Hubble limit, as we will show.

While this paper was in preparation, several papers appeared on the archive which treated related issues. The issue of multiple species and velocity bias is taken up in [17], the role of non-local galaxy bias is discussed in [18], and the importance of the equivalence principle is emphasized in [16, 18, 21]. Redshift-space and non-perturbative issues are discussed in [20]. A different kind of approximate consistency relation, one not based on symmetries in the traditional sense, is discussed in [22, 23]. Our overlap with [16] is the greatest, in that we both view the consistency relations through the prism of diffeomorphism invariance. We also overlap with [24] in the attempt to spell out the assumptions underlying the consistency relations. We are especially indebted to Filippo Vernizzi for initial collaboration and many illuminating discussions.

The paper is organized as follows. Section §2 treats the Newtonian system while §3 treats the full general relativistic problem. In §2.1, we review the symmetry discovered by KRPP and the derivation of the corresponding consistency relation, paying special attention to the underlying assumptions. In

§2.2, we discuss the robustness and limitations of the consistency relations, and delineate the underlying assumptions concerning the galaxy dynamics and biasing. We mention in §2.3 a simple Lagrangian that reproduces the Newtonian LSS equations for dark matter, though this Lagrangian is not needed for the analyses in this paper. In §3, we place the KRPP consistency relation in a larger context: all consistency relations result from residual coordinate transformations (diffeomorphisms) allowed within a given gauge. We establish the connection between such symmetries in the ζ gauge and the Newtonian gauge. The KRPP symmetry is the sub-Hubble limit of one (or more precisely, two) of these symmetries. We conclude in §4. Readers not interested in the detailed arguments can find here a summary of our main results. A few technical results are collected in the Appendices.

A few words are in order on our notation and terminology. We use the symbol π to represent the Nambu-Goldstone boson of a non-linearly realized symmetry (the pion), in accordance with standard practice. For our LSS application, π is the velocity potential. The same symbol is also used to denote the numerical value 3.14.... Which is meant should be obvious from the context. Essentially, the numerical π always precedes the Newtonian constant G in the combination $4\pi G$. In cases where both could potentially appear, we use $M_P^2 \equiv 1/(8\pi G)$ to avoid confusion. Also, we use the term *nonlinear* to refer to quantities that are not linear in the LSS observables (fields such as density or velocity). Sometimes, this has the usual meaning that such quantities go like the fields raised to higher powers: quadratic and so on. But, sometimes, this means the quantities of interest do not depend on the field variables at all, such as the nonlinear part of certain symmetry transformations e.g. Eq. (6). We rely on the context to differentiate between the two.

2 Consistency Relation from Time-dependent Translation – a Newtonian Symmetry

In this section, we focus on the Newtonian symmetry uncovered by KRPP. §2.1 is a review of the symmetry and its implied consistency relation. In §2.2 we discuss the robustness and limitations of the consistency relation, and go over what assumptions can or cannot be relaxed, especially concerning the nonlinear, astrophysically messy, galaxy observables. We discuss what kind of galaxy dynamics, and what sort of galaxy selection, could lead to violations of the consistency relation. As a by-product of our investigation, we describe a simple Lagrangian for LSS in §2.3.

2.1 Time-dependent Translation Symmetry and the Background Wave Argument – a Review

We begin with a review of the Newtonian symmetry discovered by KRPP. We go over the background wave derivation of the consistency relation in some detail, emphasizing the underlying assumptions, and making the derivation easily generalizable to the general relativistic case. Two fundamental concepts are: (1) the existence of a non-linearly realized symmetry (one that shifts at least some of the LSS observables by an amount that is independent of the observables), and (2) an adiabatic mode condition, which is an additional condition that dictates the time-dependence of the symmetry.

Time-dependent Translation Symmetry. The set of Newtonian equations of motion for LSS is: ¹

$$\delta' + \partial_i[(1 + \delta)v^i] = 0 \quad , \quad v^{i'} + v^j \partial_j v^i + \mathcal{H}v^i = -\partial_i \Phi \quad , \quad \nabla^2 \Phi = 4\pi G \bar{\rho} a^2 \delta, \quad (9)$$

where δ is the mass overdensity, $' \equiv \partial/\partial\eta$ denotes the derivative with respect to conformal time η , ∂_i denotes the derivative with respect to the comoving coordinate x^i , v^i is the peculiar velocity $dx^i/d\eta$, Φ is the gravitational potential, G is Newton's constant, $\bar{\rho}$ is the mean mass density at the time of interest, a is the scale factor, and $\mathcal{H} \equiv a'/a$ is the comoving Hubble parameter. The first equation expresses continuity or mass conservation. The second equation is the Euler equation or momentum conservation for a pressureless fluid. The third equation is the Poisson equation. Let us start with this basic set. We will later consider generalizations to include pressure, relativistic corrections, and even complex galaxy formation processes.

KRPP pointed out that this system of equations admits the following symmetry:

$$\eta \rightarrow \tilde{\eta} = \eta \quad , \quad x^i \rightarrow \tilde{x}^i = x^i + n^i \quad , \quad v^i \rightarrow \tilde{v}^i = v^i + n^{i'} \quad , \quad \Phi \rightarrow \tilde{\Phi} = \Phi - (\mathcal{H}n^{i'} + n^{i''})x^i \quad , \quad \delta \rightarrow \tilde{\delta} = \delta, \quad (10)$$

where n^i is a function of time but not space. It can be shown that under this set of transformations, Eq. (9) takes on exactly the same form, with all the variables replaced by ones with a \sim on top. To see that this is true, it is important to keep in mind:

$$\frac{\partial}{\partial \tilde{\eta}} = -n^{i'} \frac{\partial}{\partial x^i} + \frac{\partial}{\partial \eta}, \quad (11)$$

where on the left, \tilde{x}^i is held fixed, and on the right, $\partial/\partial x^i$ is at a fixed η , and $\partial/\partial \eta$ is at a fixed x^i . On the other hand $\partial/\partial \tilde{x}^i = \partial/\partial x^i$. The symmetry transformation described by Eq. (10) is a time-dependent spatial translation. (Henceforth, we will occasionally refer to this somewhat sloppily as simply translation.) Under this translation, the velocity gets shifted in the expected manner. The gravitational potential needs to be shifted correspondingly to preserve the form of the Euler equation. The density δ , on the other hand, does not change at all, in the sense that $\tilde{\delta}(\tilde{x}) = \delta(x)$.

Eq. (10) is a symmetry regardless of the time-dependence of n^i . For the purpose of deriving the consistency relations, however, we need to impose an additional condition. Suppose we start with $v^i = 0$ in Eq. (10); we would like the velocity generated by the transformation, i.e. $\tilde{v}^i = n^{i'}$, to be the long wavelength limit of an actual physical mode that satisfies the equations of motion. We say long wavelength because $n^{i'}$ has no spatial dependence and so is strictly speaking a $q = 0$ mode (q being the wavenumber/momentum in Fourier space). What we want to impose is this: if we take a physical velocity mode at a finite q , and make its q smaller and smaller, we would like $n^{i'}$ to match its time-dependence. Following terminology used in general relativity, we refer to it as the *adiabatic mode condition* [25]. This condition ensures the symmetry transformation generates a velocity mode that evolves in a physical way. It is easy to see that at long wavelength, where the equations can be linearized, Eq. (9) can be combined into a single equation:

$$\partial_i v^{i'} + \mathcal{H} \partial_i v^i = 4\pi G \bar{\rho} a^2 \partial_i \int d\eta v^i, \quad (12)$$

¹Throughout this paper, we will be cavalier about the placement of indices for objects with Latin indices, e.g. v^i, x^i are the same as v_i, x_i .

or written in a more familiar form:

$$\delta'' + \mathcal{H}\delta' = 4\pi G\bar{\rho}a^2\delta \quad \text{with} \quad \delta' = -\partial_i v^i. \quad (13)$$

One might be tempted to say a velocity going like $n^{i'}$ satisfies Eq. (12) trivially for an $n^{i'}$ of arbitrary time-dependence, since $n^{i'}$ has no spatial dependence. What we want, however, is for $n^{i'}$ to have the same time-dependence as that of a velocity mode at a low, but finite momentum. In other words, we impose the *adiabatic mode condition*:

$$n^{i''} + \mathcal{H}n^{i'} = 4\pi G\bar{\rho}a^2 n^i \quad (14)$$

i.e. $n^i(\eta)$ has the same time-dependence as the linear growth factor $D(\eta)$, assuming growing mode initial conditions. Effectively, we demand that our symmetry-generated velocity-shift (or more precisely, the nonlinear part thereof) satisfy Eq. (12) with the spatial gradient removed.

The Background Wave Argument. Next, we give the background wave derivation of the consistency relation. More sophisticated and rigorous derivations exist [7, 8, 11–13], but the background wave argument has the virtue of being fairly intuitive. Our goal here is to go over the underlying assumptions, and formulate the argument in such a way to ease later generalizations. The form of our expressions follow closely those in H2K [7].

Before we carry out the argument, it is convenient (especially for later discussions) to introduce the velocity potential π , assuming potential flow on large scales,² in which case, the symmetry transformation of Eq. (10) tells us:

$$\pi \rightarrow \tilde{\pi} = \pi + n^{i'}x^i \quad \text{with} \quad v^i = \partial_i \pi. \quad (15)$$

The velocity potential π has the hallmark of a pion or Nambu-Goldstone boson: it experiences a *nonlinear* shift under the symmetry transformation.³ Note that π also implicitly has a *linear* shift:

$$\tilde{\pi}(\tilde{x}) = \pi(x) + n^{i'}x^i \sim \pi(\tilde{x}) - n^i \partial_i \pi + n^{i'}x^i \quad (16)$$

where we have Taylor expanded $\pi(x)$ to first order in $\tilde{x}^i - x^i$, assuming a small n^i . Here, linear and nonlinear refer to whether or not the transformation is linear in π (or any other LSS fields). The total shift in π , i.e. $\tilde{\pi}(\tilde{x}) - \pi(\tilde{x})$, thus has both linear and nonlinear pieces:

$$\Delta_{\text{lin.}}\pi = -n^i \partial_i \pi \quad , \quad \Delta_{\text{nl.}}\pi = n^{i'}x^i. \quad (17)$$

Similarly, the overdensity changes by the amount $\tilde{\delta}(\tilde{x}) - \delta(\tilde{x})$:

$$\Delta_{\text{lin.}}\delta = -n^i \partial_i \delta \quad , \quad \Delta_{\text{nl.}}\delta = 0, \quad (18)$$

² Assuming potential flow is not strictly necessary, as the argument can be made using the velocity v^i itself in place of $\partial_i \pi$. The reason we make this assumption is that the velocity enters into our derivation mainly as a large scale or low momentum mode. Assuming the growing mode initial condition, the large scale velocity does take the form of a potential flow. Indeed, vorticity remains zero until orbit crossing. We do *not* assume potential flow on small scales.

³Note that Φ also experiences a nonlinear shift (Eq. 10), and thus can also be used as the pion in this derivation. The two give the same result, see §3.3.

i.e. δ experiences no nonlinear shift.

Consider an N-point function involving a product of N LSS observables. Let us denote each as \mathcal{O} , labeled by momentum, so that the N-point function is $\langle \mathcal{O}_{\vec{k}_1} \mathcal{O}_{\vec{k}_2} \dots \mathcal{O}_{\vec{k}_N} \rangle$. Here, the \mathcal{O} 's at different momenta need not be the same observable. For instance, one can be δ , the other can be the gravitational potential, et cetera. They need not even be evaluated at the same time. We are interested in this N-point function in the presence of some long wavelength (soft) π . Let us imagine splitting all fluctuations into hard and soft modes, with $\vec{k}_1, \dots, \vec{k}_N$ falling into the hard category. The N-point function obtained by integrating over the hard modes, but leaving the soft modes of π unintegrated, can be Taylor expanded as:

$$\langle \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle_{\pi_{\text{soft}}} \approx \langle \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle_0 + \int \frac{d^3 \vec{p}}{(2\pi)^3} \frac{\delta \langle \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle_{\pi_{\text{soft}}}}{\delta \pi_{\vec{p}}^*} \Big|_0 \pi_{\vec{p}}^*, \quad (19)$$

where we have taken the functional derivative with respect to, and summed over the Fourier modes of, π with *soft* momenta \vec{p} . Multiplying both sides by $\pi_{\vec{q}}$ (where \vec{q} is also soft) and ensemble averaging over the soft modes, one finds

$$\frac{\langle \pi_{\vec{q}} \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle}{P_{\pi}(q)} = \frac{\delta \langle \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle_{\pi_{\text{soft}}}}{\delta \pi_{\vec{q}}^*} \Big|_0. \quad (20)$$

We have used the definition of the power spectrum: $\langle \pi_{\vec{q}} \pi_{\vec{p}}^* \rangle = (2\pi)^3 \delta_D(\vec{q} - \vec{p}) P_{\pi}(q)$, with δ_D being the Dirac delta function. We can on the other hand compute the derivative on the right hand side this way:

$$\int \frac{d^3 \vec{p}}{(2\pi)^3} \frac{\delta \langle \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle_{\pi_{\text{soft}}}}{\delta \pi_{\vec{p}}^*} \Big|_0 \Delta_{\text{nl.}} \pi_{\vec{p}}^* = \Delta_{\text{lin.}} \langle \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle. \quad (21)$$

This statement says that the change to the N-point function induced by the symmetry transformation (the right hand side) is equivalent to the change to the N-point function by adding a long-wavelength background π induced by the same symmetry (the left hand side). We will unpack it a bit more in §2.2. A careful reader might note that there is no reason why one should include on the right hand side only the *linear* part of the transformation of the N-point function. That is true: by including only the linear transformation, we are effectively dealing with the *connected* N-point function. For a proof, see H2K.⁴ Combining Eqs. (20) and (21), and adding the superscript *c* for connected N-point function, we have:

$$\boxed{\int \frac{d^3 q}{(2\pi)^3} \frac{\langle \pi_{\vec{q}} \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^c}{P_{\pi}(q)} \Delta_{\text{nl.}} \pi_{\vec{q}}^* = \Delta_{\text{lin.}} \langle \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^c.} \quad (22)$$

It is important to note that the connected correlation functions on both sides contain delta functions. This way of writing the consistency relation follows the Ward identity treatment of H2K, and is applicable

⁴The restriction to the *connected* N-point function will not be so relevant for the Newtonian (KRPP) consistency relation or the general relativistic dilation consistency relation, i.e. they take the same form whether the consistency relation is phrased in terms of *connected* or *general* correlation functions. The restriction is relevant for those consistency relations that involve more than one derivative on the right hand side. For them, it is important to keep in mind that the N momenta on the right hand side sum to zero (see [11]); deriving *general* consistency relations from *connected* ones necessarily introduce extra terms.

to any symmetries with π as the Nambu-Goldstone boson. The background wave argument has the advantage of being intuitive, but is a bit heuristic. Readers interested in subtleties can consult e.g. H2K. Our final result here matches theirs.

Let us apply Eq. (22) to the translation symmetry. To be specific, let us take our observable \mathcal{O} to be the mass overdensity δ . We use the following convention for the Fourier transform of some function f :

$$f(\vec{q}) = \int d^3x f(\vec{x}) e^{i\vec{q}\cdot\vec{x}}, \quad f(\vec{x}) = \int \frac{d^3q}{(2\pi)^3} f(\vec{q}) e^{-i\vec{q}\cdot\vec{x}}. \quad (23)$$

Eqs. (17) and (18) thus implies:

$$\Delta_{\text{nl.}} \pi_{\vec{q}}^* = i n^{j'} \frac{\partial}{\partial q^j} [(2\pi)^3 \delta_D(\vec{q})] \quad , \quad \Delta_{\text{lin.}} \delta_{\vec{k}} = i n^j k^j \delta_{\vec{k}}. \quad (24)$$

Substituting into Eq. (22), we find

$$\lim_{\vec{q} \rightarrow 0} n^{j'}(\eta) \frac{\partial}{\partial q^j} \left[\frac{\langle \pi_{\vec{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^c}{P_\pi(q)} \right] = - \sum_{a=1}^N n^j(\eta_a) k_a^j \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^c, \quad (25)$$

where η is the time implicitly assumed for $\pi_{\vec{q}}$, and η_a is the time for $\delta_{\vec{k}_a}$. Note that k_a^j refers to the j -component of the vector \vec{k}_a . Since $n^j(\eta)$ has the time-dependence of the linear growth factor $D(\eta)$ (from the adiabatic mode condition), but can otherwise point in an arbitrary direction, we conclude:

$$\lim_{\vec{q} \rightarrow 0} \frac{\partial}{\partial q^j} \left[\frac{\langle \pi_{\vec{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^c}{P_\pi(q)} \right] = - \sum_{a=1}^N \frac{D(\eta_a)}{D'(\eta)} k_a^j \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^c. \quad (26)$$

We have yet to remove the delta functions from both sides. To do so, we use the following, pure shift, symmetry:

$$\pi \rightarrow \tilde{\pi} = \pi + \text{const.}, \quad (27)$$

This symmetry does not involve transforming space-time at all, and so none of the observables receive a *linear* shift. The argument leading to Eq. (22) thus tells us

$$\lim_{\vec{q} \rightarrow 0} \frac{\langle \pi_{\vec{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}}{P_\pi(q)} = 0. \quad (28)$$

We use the superscript c' to denote the connected correlation function with the overall delta function removed:

$$\langle \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^c = (2\pi)^3 \delta_D(\vec{k}_1 + \dots + \vec{k}_N) \langle \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^{c'} \quad (29)$$

Combining Eqs. (26) and (28), we have the Newtonian *translation consistency relation*:

$$\boxed{\lim_{\vec{q} \rightarrow 0} \frac{\partial}{\partial q^j} \left[\frac{\langle \pi_{\vec{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}}{P_\pi(q)} \right] = - \sum_{a=1}^N \frac{D(\eta_a)}{D'(\eta)} k_a^j \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'},} \quad (30)$$

where η is the time for the soft mode $\pi_{\vec{q}}$, and each η_a is the time for the corresponding hard mode $\delta_{\vec{k}_a}$. This turns out to be a common feature for all consistency relations as we will see: Eqs. (28) and (26) are two consistency relations that differ by one derivative with respect to q ; the former allows us to remove the delta function from the latter in a straightforward way – the result is Eq. (30). The form adopted by KRPP is to integrate the above over q , and using Eq. (28), to obtain:⁵

$$\lim_{\vec{q} \rightarrow 0} \left[\frac{\langle \pi_{\vec{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}}{P_\pi(q)} \right] = - \sum_{a=1}^N \frac{D(\eta_a)}{D'(\eta)} \vec{q} \cdot \vec{k}_a \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}. \quad (31)$$

The soft mode $\pi_{\vec{q}}$, by the linearized continuity equation, is related to $\delta_{\vec{q}}$ by⁶

$$\delta_{\vec{q}}(\eta) = q^2 \frac{D(\eta)}{D'(\eta)} \pi_{\vec{q}}(\eta). \quad (32)$$

One can therefore rewrite Eq. (31) as⁷

$$\lim_{\vec{q} \rightarrow 0} \left[\frac{\langle \delta_{\vec{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}}{P_\delta(q)} \right] = - \sum_{a=1}^N \frac{D(\eta_a)}{D(\eta)} \frac{\vec{q} \cdot \vec{k}_a}{q^2} \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}. \quad (33)$$

In the context of this Newtonian derivation, we would like to think of the form expressed in Eq. (30) and Eq. (28) as more fundamental, since it is π that experiences a nonlinear shift in the symmetry transformation, acting as the pion, and since the expression in Eq. (33) contains two non-relativistic consistency relations, one trivial and one nontrivial. It is also worth stressing that Eq. (30) does not constrain $\partial^2/\partial q^2$ of $[\langle \pi_{\vec{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}/P_\pi(q)]$ in the soft limit, i.e. Eq. (31) can in principle contain $O(q^2)$ corrections, and Eq. (33) can contain $O(q^0)$ corrections.

Let us close this derivation by observing that the only assumption made about the hard modes is how they transform under the symmetry (in particular, the linear part of their transformation; see Eqs. 18, 24). Thus, suppose we have some observable \mathcal{O} whose linear transformation under the spatial translation is:

$$\Delta_{\text{lin.}} \mathcal{O}_{\vec{k}} = i n^j g^j \mathcal{O}_{\vec{k}}. \quad (34)$$

For instance, if \mathcal{O} is the galaxy overdensity, we expect $g^j = k^j$, but g^j could take other forms for other observables. Exactly the same derivation then gives the Newtonian translation consistency relation in a more general form:

$$\lim_{\vec{q} \rightarrow 0} \frac{\partial}{\partial q^j} \left[\frac{\langle \pi_{\vec{q}} \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^{c'}}{P_\pi(q)} \right] = - \sum_{a=1}^N \frac{D(\eta_a)}{D'(\eta)} g_a^j \langle \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^{c'}, \quad (35)$$

⁵ We are grateful to Lasha Berezhiani and Justin Khoury for helping us understand this point.

⁶ In the context of a consistency relation which is purported to be non-perturbative, one might wonder if using the linear relation between δ and π (for the soft mode only) is justified. It can be shown that including nonlinear corrections to this relation leads to terms subdominant in the squeezed limit of the correlation function. See [20].

⁷ The use of the equation of motion within the *connected* correlation function does not lead to contact terms. See e.g. H2K and [11] on the role of contact terms in Ward identity arguments.

where we have allowed the possibility that the N hard modes correspond to different observables, thus potentially a different g_a^j for each $a = 1, \dots, N$. Corollaries – analogs of Eqs. 31, 33 – follow in the same way:

$$\lim_{\vec{q} \rightarrow 0} \left[\frac{\langle \pi_{\vec{q}} \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^{c'}}{P_\pi(q)} \right] = - \sum_{a=1}^N \frac{D(\eta_a)}{D'(\eta)} \vec{q} \cdot \vec{g}_a \langle \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^{c'}, \quad (36)$$

and

$$\lim_{\vec{q} \rightarrow 0} \left[\frac{\langle \delta_{\vec{q}} \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^{c'}}{P_\pi(q)} \right] = - \sum_{a=1}^N \frac{D(\eta_a)}{D(\eta)} \frac{\vec{q} \cdot \vec{g}_a}{q^2} \langle \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^{c'}, \quad (37)$$

where it should be understood that Eq. (36) and Eq. (37) contain $O(q^2)$ and $O(q^0)$ corrections respectively. Phrased as such, the consistency relation is fairly robust: the detailed dynamics of the hard modes has no relevance; it matters not whether the corresponding observables are astrophysically messy or highly nonlinear. All we need to know is how they transform under a spatial translation. To understand this robustness better, it is helpful to study concrete examples, which is the subject of the next section.

2.2 Robustness and Limitations of the Consistency Relation(s)

To understand better the robustness of the consistency relation, it is instructive to ask the question: when does it fail? As we will see, the consistency relation stands on three legs: the existence of the time-dependent translation symmetry, the single-field initial condition, and the adiabatic mode condition. All three are necessary in order for the consistency relation to hold. Here, we focus on the KRPP consistency relation as a specific example, but the points we raise are general, pertaining to other consistency relations (§3) as well.

1. Initial condition: the single-field assumption. A crucial step in the derivation is Eq. (21): that linear transformations of a collection of hard modes, represented by $\mathcal{O}_{\vec{k}}$, can be considered equivalent to placing the same hard modes in the presence of a soft mode – the pion $\pi_{\vec{q}}$. That this single soft mode is sufficient to account for all the transformations of the hard modes is an assumption about initial conditions. In the context of inflation, the assumption is often phrased as that of a single field or a single clock. In our Newtonian LSS context, in addition to keeping only the growing modes, essentially the assumption is that of Gaussian initial conditions.⁸ More precisely, one demands that the initial condition does not contain a coupling between soft and hard modes beyond that captured by Eq. (21). We will follow the inflation terminology and call this the *single-field assumption*.

2. Adiabatic mode condition. Another crucial ingredient in the derivation is the *adiabatic mode condition*, that the symmetry transformation have the correct time-dependence so that the nonlinear shift $\Delta_{\text{nl}}\pi$ may be the long wavelength limit of an actual physical mode. We stress that this is an additional requirement on top of demanding a symmetry, with some non-trivial implications. To spell them out, it is useful to have a concrete example. Since the consistency relation is purported to be robust, in the sense that the hard modes can be highly nonlinear and even astrophysically complex, let us write

⁸Note that in the inflationary context, Gaussianity is also assumed – for the wave-function in the far past.

down a system of equations that allow for these complexities:

$$\delta'_{(a)} + \vec{\nabla} \cdot [(1 + \delta_{(a)})\vec{v}_{(a)}] = R_{(a)} \quad , \quad \vec{v}_{(a)}' + \vec{v}_{(a)} \cdot \vec{\nabla} \vec{v}_{(a)} + \mathcal{H}\vec{v}_{(a)} = -\vec{\nabla} \Phi + \vec{F}_{(a)} . \quad (38)$$

Here a labels the species: for instance, it can be dark matter, populations of galaxies, baryons and so on. $R_{(a)}$ represents a source term for the density evolution. For dark matter, we expect $R_{(a)} = 0$ (barring significant annihilation or decay). For galaxies, $R_{(a)}$ quantifies the effect of galaxy formation and mergers. All particles are subjected to the same gravitational force plus a species-dependent force $\vec{F}_{(a)}$. The gravitational potential Φ is sourced by the total mass fluctuation:

$$\nabla^2 \Phi = 4\pi G a^2 \bar{\rho} \delta_T , \quad (39)$$

where δ_T represents the effective total mass density fluctuation from all particles.

A natural generalization of the (time-dependent) translational symmetry from the previous section would be:

$$\begin{aligned} \eta &\rightarrow \tilde{\eta} = \eta \quad , \quad x^i \rightarrow \tilde{x}^i = x^i + n^i \quad , \quad v_{(a)}^i \rightarrow \tilde{v}_{(a)}^i = v_{(a)}^i + n^{i'} \quad , \quad \Phi \rightarrow \tilde{\Phi} = \Phi - (\mathcal{H}n^{i'} + n^{i''})x^i \quad , \\ \delta_{(a)} &\rightarrow \tilde{\delta}_{(a)} = \delta_{(a)} \quad , \quad \delta_T \rightarrow \tilde{\delta}_T = \delta_T \quad , \quad R_{(a)} \rightarrow \tilde{R}_{(a)} = R_{(a)} \quad , \quad F_{(a)}^i \rightarrow \tilde{F}_{(a)}^i = F_{(a)}^i \end{aligned} \quad (40)$$

This means that $R_{(a)}$ and $\vec{F}_{(a)}$ remain invariant under this symmetry. For instance, they are invariant if R and \vec{F} depend only on δ , on the spatial gradient of \vec{v} , or on the second gradients of Φ . R and \vec{F} could even have explicit dependence on time η . What would violate invariance is if R or \vec{F} depends on \vec{v} with no gradients – that is, unless the dependence is of the form $\vec{v}_{(a)} - \vec{v}_{(b)}$ (in which case the shifts in the velocities of the two different species (a) and (b) cancel out). Specializing to the case of galaxies, what this means is that the number density evolution and dynamics of the galaxies does not care about the absolute size of the velocity, but only about the velocity difference (either between neighbors, or between species). The only context in which the absolute size of velocity plays a role is through Hubble friction – this is the origin of the \mathcal{H} dependent term in the nonlinear shift of Φ . In other words, Hubble friction aside, *galaxy formation and dynamics is frame invariant*, which seems a fairly safe assumption. For instance, dynamical friction or ram pressure, which no doubt exerts an influence on galaxies, should depend on velocity difference.

Thus, let us assume Eq. (40) is a symmetry of our system – note that this is a symmetry regardless of the time-dependence of n^i . As emphasized earlier, this is not enough to guarantee the validity of the consistency relation. To derive the consistency relation, n^i must have the correct time-dependence: $n^{i'}$ (the nonlinear shift in v^i) must match the time-dependence of a physical long-wavelength velocity perturbation.⁹ This has to hold for all species, meaning the *same* $n^{i'}$ matches the long-wavelength velocity perturbation of each and everyone of the species. In other words, *all species should move with the same velocity on large scales*. This leads to two subtleties, which are best illustrated by assuming an explicit form for \vec{F} . Consider:

$$\vec{F}_{(a)} = -c_{s(a)}^2 \vec{\nabla} \ln(1 + \delta_{(a)}) - \beta_{(a,b)}(\vec{v}_{(a)} - \vec{v}_{(b)}) - \alpha_{(a)} \vec{\nabla} \varphi \quad , \quad \nabla^2 \varphi = 8\pi G \sum_a \alpha_{(a)} a^2 \bar{\rho}_a \delta_a . \quad (41)$$

⁹ This typically means n^i should satisfy Eq. (14), i.e. n^i must evolve like the linear growth factor D . This holds if all particles, on large scales, evolve like dark matter and if gravity is the only long-range force. See further discussion below.

The first term on the right of the expression for \vec{F} represents some sort of pressure – c_s is the sound speed – this would be relevant if the subscript (a) represents baryons at finite temperature. The second term represents some sort of friction that depends on the velocity difference between two species, with a coefficient β . The third term represents an additional fifth force, mediated by the scalar φ , with a coupling α . The scalar φ obeys a Poisson-like equation. In scalar-tensor theories, the tensor part of the theory mediates a universal gravitational force (described by the gravitational potential Φ), but the scalar need not be universally coupled: hence we allow the coupling $\alpha_{(a)}$ to depend on the species (see e.g. [26, 27]). The form for the additional force \vec{F} proposed in Eq. (41) is fairly generic: counting derivatives, we can see that the pressure term goes like $\partial\delta$, whereas the other two terms go like $\partial^{-1}\delta$. In terms of the symmetry transformation, one can see that

$$\varphi \rightarrow \tilde{\varphi} = \varphi \quad (42)$$

is compatible with Eq. (40). Thus, we have a system (Eq. 38) that respects the translational symmetry spelled out in Eq. (40), even if many different kinds of forces are present, including non-gravitational or modified gravitational ones such as in Eq. (41). We wish to see how, despite the presence of the (time-dependent) translational symmetry, there can still be a breakdown of the consistency relation, due to obstructions in satisfying the adiabatic mode condition – that the velocity perturbations of all species should be equal on large scales.

2a. Soft dynamics constraint. In the long wavelength limit, one can ignore the pressure term compared to the other two terms in the expression for $\vec{F}_{(a)}$. Let us first focus on the fifth-force term. This term is at the same level in derivative as the normal gravitational force $-\vec{\nabla}\Phi$, and thus both have to be taken into account on large scales. The problem with a long-range fifth-force is the non-universal coupling: if there is a different coupling $\alpha_{(a)}$ for each kind of particles, the different species will move with different velocities even on large scales. This means no *single* n^{ii} can possibly generate long-wavelength velocity perturbations for all species. In other words, unless the soft (large-scale) dynamics obeys the equivalence principle, the consistency relation would be violated, as emphasized by [18, 21]. We stress that, in our example, the violation of equivalence principle occurs *without* the violation of the translation symmetry described by Eq. (40). The fact that the consistency relation is not obeyed is entirely because of the failure to satisfy the adiabatic mode condition when the equivalence principle is violated. The friction term (second term on the right of Eq. 41), on the other hand, is compatible with the adiabatic mode condition – it simply vanishes if the velocities of different species are equal, and is therefore consistent with the large scale requirement that all species flow with the same velocity. To sum up, the *soft dynamics constraint* is: for the consistency relation to be valid, the dynamics on large scales must be consistent with all species moving with the same velocity.

2b. Squeezing constraint. Let us next turn to the pressure term (first term on the right of Eq. 41). Since different species have different sound speeds, this also leads to differences in velocity flows. This is relatively harmless though, since the pressure term becomes subdominant on large scales. Thus, there is no problem with the adiabatic mode condition, which is really a condition on motions in the soft limit $q \rightarrow 0$. The presence of pressure does lead to a practical limitation on the application of the consistency relation, however. The consistency relation is a statement about an $(N+1)$ -point function in the squeezed limit $q \ll k_1, \dots, k_N$. There is the practical question of how small q has to be. An important requirement is: q must be sufficiently small such that the velocity perturbations of different species have the same

time-dependence as that generated by a single $n^{i'}$. In the present context, it means $q < \mathcal{H}/c_s$, i.e. the length scale must be above the Jeans scale.¹⁰ We refer to this as the *squeezing constraint*: the soft leg of the consistency relation must be sufficiently soft that any difference in force on the different species becomes negligible. This is worth emphasizing, because clearly dark matter and baryons are subject to different forces: while that does not by itself lead to the breakdown of the consistency relations, one has to be careful to make sure that the squeezed correlation function is sufficiently squeezed.

3. Galaxy-biasing. It is also instructive to approach the subject of consistency relation violation from the viewpoint of galaxy-biasing. What kind of galaxy-biasing would lead to the violation of consistency relation? We can only address this question in the perturbative regime, but it nonetheless provides some useful insights. Suppose the galaxy overdensity $\delta_{(a)}$ (of type a) and matter density δ are related by:

$$\delta_{(a)\vec{k}} = b_{(a)} \left(\delta_{\vec{k}} + \int \frac{d^3 k'}{(2\pi)^3} \frac{d^3 k''}{(2\pi)^3} (2\pi)^3 \delta_D(\vec{k} - \vec{k}' - \vec{k}'') W_{(a)}(\vec{k}', \vec{k}'') [\delta_{\vec{k}'} \delta_{\vec{k}''} - \langle \delta_{\vec{k}'} \delta_{\vec{k}''} \rangle] \right), \quad (43)$$

where $b_{(a)}$ is a linear bias factor (independent of momentum) and $W_{(a)}$ is a kernel that describes a general quadratic bias. To the lowest order in perturbation theory, it can be shown that the bispectrum between three types of galaxies a , b and c , at momenta $\vec{q}, \vec{k}_1, \vec{k}_2$ and times η, η_1, η_2 respectively, is

$$\begin{aligned} B_{(abc)}(\vec{q}, \eta; \vec{k}_1, \eta_1; \vec{k}_2, \eta_2) = & \\ & 2b_{(a)}b_{(b)}b_{(c)}P_\delta(k_1, \eta, \eta_1)P_\delta(k_2, \eta, \eta_2) \left[\left(\frac{5}{7} + \frac{1}{2}\hat{k}_1 \cdot \hat{k}_2 \left(\frac{k_1}{k_2} + \frac{k_2}{k_1} \right) + \frac{2}{7}(\hat{k}_1 \cdot \hat{k}_2)^2 \right) + W_{(a)}(-\vec{k}_1, -\vec{k}_2) \right] \\ & + 2b_{(a)}b_{(b)}b_{(c)}P_\delta(q, \eta, \eta_1)P_\delta(k_2, \eta_1, \eta_2) \left[\left(\frac{5}{7} + \frac{1}{2}\hat{q} \cdot \hat{k}_2 \left(\frac{q}{k_2} + \frac{k_2}{q} \right) + \frac{2}{7}(\hat{q} \cdot \hat{k}_2)^2 \right) + W_{(b)}(-\vec{q}, -\vec{k}_2) \right] \\ & + 2b_{(a)}b_{(b)}b_{(c)}P_\delta(q, \eta, \eta_2)P_\delta(k_1, \eta_1, \eta_2) \left[\left(\frac{5}{7} + \frac{1}{2}\hat{q} \cdot \hat{k}_1 \left(\frac{q}{k_1} + \frac{k_1}{q} \right) + \frac{2}{7}(\hat{q} \cdot \hat{k}_1)^2 \right) + W_{(c)}(-\vec{q}, -\vec{k}_1) \right] \end{aligned} \quad (44)$$

where P_δ is the linear mass power spectrum – its two time-arguments signify the fact that the two δ 's involved can be at different times. We are interested in the $\vec{q} \rightarrow 0$ limit:

$$\begin{aligned} & \lim_{\vec{q} \rightarrow 0} b_{(a)} \frac{B_{(abc)}(\vec{q}, \eta; \vec{k}_1, \eta_1; \vec{k}_2, \eta_2)}{P_{(aa)}(q, \eta, \eta)} \\ &= E + \frac{D(\eta_1)}{D(\eta)} \left[\frac{\vec{q} \cdot \vec{k}_2}{q^2} + 2W_{(b)}(-\vec{q}, -\vec{k}_2) + O(q^0) \right] P_{(bc)}(k_2, \eta_1, \eta_2) \\ & \quad + \frac{D(\eta_2)}{D(\eta)} \left[\frac{\vec{q} \cdot \vec{k}_1}{q^2} + 2W_{(c)}(-\vec{q}, -\vec{k}_1) + O(q^0) \right] P_{(bc)}(k_1, \eta_1, \eta_2) \\ &= E - \left[\frac{D(\eta_1)}{D(\eta)} \left(\frac{\vec{q} \cdot \vec{k}_1}{q^2} - 2W_{(b)}(-\vec{q}, \vec{k}_1) \right) + \frac{D(\eta_2)}{D(\eta)} \left(\frac{\vec{q} \cdot \vec{k}_2}{q^2} - 2W_{(c)}(-\vec{q}, \vec{k}_2) \right) + O(q^0) \right] P_{(bc)}(k_1, \eta_1, \eta_2), \\ & \text{where } E \equiv 2 \frac{D(\eta)^2}{D(\eta_1)D(\eta_2)} \frac{b_{(a)}^2}{b_{(b)}b_{(c)}} \frac{P_{(bc)}(k_1, \eta_1, \eta_2)P_{(bc)}(k_2, \eta_1, \eta_2)}{P_{(aa)}(q, \eta, \eta)} \left[O(q^2) + W_{(a)}(-\vec{k}_1, -\vec{k}_2) \right]. \end{aligned} \quad (45)$$

We have used $P_{(aa)}(q, \eta, \eta) = (D(\eta)/D(\eta_1))b_{(a)}^2 P_\delta(q, \eta, \eta_1)$, $P_{(bc)}(k_1, \eta_1, \eta_2) = b_{(b)}b_{(c)}P_\delta(k_1, \eta_1, \eta_2)$ and so on (appropriate only perturbatively). Comparing this expression with the consistency relation expressed

in Eq. (37) (identifying \vec{g}_a with \vec{k}_a), we see that the two agree if E , $W_{(b)}$ and $W_{(c)}$ can be ignored, and $b_{(a)} = 1$. A number of comments are in order.

First, let us focus on the case with no galaxy biasing, so that Eq. (45) simply constitutes a perturbative check of the consistency relation for the mass overdensity i.e. Eq. (33). We see that the term E can be ignored compared to the terms kept only if the soft power spectrum is not too blue: assuming $P_{(aa)}(q) \sim q^n$ for small q , the validity of the consistency relation requires $n < 3$. This is a limitation on the consistency relation that is not often emphasized. In practice though, the realistic power spectrum has no problem satisfying this requirement.

Let us next consider the effects of galaxy biasing. The second point we would like to raise is that the soft mode must be kept unbiased. There are two reasons for this, one trivial, the other less so. The trivial reason is that the left hand side of the consistency relation has to be corrected by a factor of $b_{(a)}$, the linear bias factor for the soft mode. This is not a big problem: one can obtain an estimate of the linear bias and correct the consistency relation when comparing against observations of the galaxy bispectrum. The more non-trivial problem is the presence of the quadratic bias kernel $W_{(a)}$ in E . Consider for instance a local biasing model of the form: $\delta_{(a)} = b_{(a)}\delta + b_{(a),2}\delta^2/2$ in real space, where $b_{(a)}$ and $b_{(a),2}$ are constants, typically referred as the linear and quadratic bias factors. In this case $W_{(a)} = b_{(a),2}/(2b_{(a)})$ has no momentum dependence, and so E contains a contribution that goes like q^{-n} for $P_{(aa)} \sim q^n$. This means one needs $n < 1$ for E to be negligible compared to the terms we keep in the consistency relation. On the largest scales, n approaches 1, though observations suggest it is slightly less than 1. On smaller scales (but still keeping $q \ll k_1, \dots, k_N$), the relevant n is on the safe side. Nonetheless, this perturbative check suggests that one should be careful in using the a biased observable for the soft-mode.

Henceforth, let us assume the soft-mode is unbiased but the hard modes are biased, in which case E is safely negligible in the squeezed limit as long as $n < 3$. The third point we wish to raise is that the validity of the consistency relation requires the hard modes be biased in a way that is not too infrared-divergent: $W_{(b)}(-\vec{q}, \vec{k}_1)$ and $W_{(c)}(-\vec{q}, \vec{k}_2)$ cannot contain terms that go like k_1/q or k_2/q i.e.

$$\boxed{W_{(b)}(-\vec{q}, \vec{k}), W_{(c)}(-\vec{q}, \vec{k}) < 1/q} \quad (46)$$

in the $q \rightarrow 0$ limit. As mentioned above, the local biasing model typically assumed in LSS studies implies the kernels $W_{(b)}$ and $W_{(c)}$ are momentum independent, and is thus consistent with the consistency relation. It is worth emphasizing that the word *local* in local biasing is a bit misleading: it merely states that the galaxy density at a given point in real space is related to the mass density at the same point. In reality, galaxies form out of the collapse of larger regions, influenced by the tidal field of the environment: there are therefore good reasons to believe that galaxy biasing is at some level *non-local*, i.e. the galaxy density at a given point is affected by the mass density at other points. This non-locality is not non-locality in the field theory sense, in that there is nothing non-local in the dynamics, and the so-called non-local galaxy bias arises completely out of local processes. A violation of the consistency relation requires more than

¹⁰Since the Jeans scale changes with time in general, the formal requirement is therefore that the soft-mode be longer than the Jeans scale at all times: $q < (\mathcal{H}/c_s)_{rec.}$ where the maximum size of the sound horizon may be conservatively estimated to be the size at recombination. In practice, since most of the present day non-Gaussianity is generated at late times, it is typically sufficient to require that the contribution from early times be subdominant. Using second order perturbation theory, we estimate that $q < \mathcal{H}/c_{s_{rec.}}(D(\eta_{obs.})/D(\eta_{rec.}))$ is a sufficient parametric condition for Eq. 30 to be valid (at least away from the equal time limit).

a non-local galaxy bias though. It requires the non-local biasing kernel to be infrared divergent. This does not appear to be so easily obtained in a random-walk halo-biasing model, even if tidal effects are taken into account [28]. One way it arises is in a model in which galaxies are born with a velocity bias, as pointed out by [29]: the quadratic kernel $W_{(c)}$ (or $W_{(b)}$), for some galaxy population with a velocity bias of b_v^* at birth (i.e. the galaxy velocity equals b_v^* times the dark matter velocity when the galaxy forms):

$$W_{(c)}(\vec{q}, \vec{k}) \sim 2b_{(c)}^{-1}(b_v^* - 1) \left(\frac{D_*}{D} \right)^{3/2} \hat{q} \cdot \hat{k} \left(\frac{q}{k} + \frac{k}{q} \right), \quad (47)$$

where D_* is the linear growth factor at the time of birth, and D is growth factor at the time of interest. We display only the term that has a dipolar dependence on the angle between \vec{q} and \vec{k} , and have taken the late time limit. This has precisely the kind of infrared divergence in the $q \ll k$ limit which would invalidate the consistency relation. It is interesting that this is also an example where we should have expected a violation of the consistency relation based on earlier arguments – the existence of a scale-independent velocity bias b_v^* means dark matter and galaxies do not flow in the same way, even on large scales. This violates the adiabatic mode condition, and so it is not a surprise that the consistency relation fails. Realistically, velocity bias is present at some level of course, but is expected to approach unity on sufficiently large scales, unless of course the equivalence principle is violated [21]. As a general statement, we can say that a non-local galaxy bias that is more infrared-divergent than Eq. (46) is what one needs to violate the consistency relation. It is interesting to ask whether there are other ways to physically generate such a galaxy bias besides through equivalence principle violations. This naturally brings us to the issue of selection.

It is worth emphasizing that the galaxy bias is also partly a selection bias: one chooses to study galaxies of a certain luminosity, color, morphology or some other property of interest. The question is then: can one choose the galaxy sample in such a way as to violate Eq. (46)? What if one chooses galaxies based on their motions, for instance, selecting galaxies that have systematically large velocities? It would seem by hand we have introduced a velocity bias, and thus a violation of the consistency relation. This is actually not a violation in the technical sense. Choosing galaxies based on their motions can be thought of as weighing the galaxies by velocities, i.e. $\delta_g \rightarrow \delta_g(1 + \pi)$. From the point of view of violating the consistency relation, it is most relevant to consider weighing by the large scale velocity. In that case, it is not surprising one finds additional terms that diverge in the squeezed limit of the correlation function – this is because we have included in the correlation function additional soft modes that carry with them additional powers $1/q$.¹¹

4. Robustness. Let us close this section by reiterating how robust the consistency relation is. As long as the underlying assumptions – the existence of the time-dependent translation symmetry, single field initial condition and adiabatic mode condition – are satisfied, the consistency relation is robust. The hard or high momentum modes can be those of any LSS observable (or even mixtures of observables), referred to as $\mathcal{O}_{\vec{k}}$ in Eq. (37). No assumption is made about the size of $\mathcal{O}_{\vec{k}}$: it matters not at all how nonlinear or non-perturbative these high momentum observables are. Indeed, we do not even need to know their detailed dynamics: all we need to know is how they transform under the symmetry in Eq. (34).¹² Their evolution can be a lot more complicated than that of dark matter (e.g. Eq. 38). In other words, $\mathcal{O}_{\vec{k}}$ can

¹¹We thank Paolo Creminelli and Guido D’Amico for discussions on this point.

¹²If the observable turns out to transform differently from Eq. (34), one can go back to the more fundamental Eq. (22) to figure out the correct consistency relation.

be astrophysically messy observables, such as those associated with galaxies. The presence of pressure effects, multiple components, multiple-streaming,¹³ star formation, supernova explosions, etc. does not lead to violations of the consistency relations, as long as the adiabatic mode condition – i.e. the soft dynamics constraint and the squeezing constraint – is satisfied. This is why the LSS consistency relations are interesting: they provide a reliable window into the non-perturbative, astrophysically complex regime.

2.3 A Simple Fluid Lagrangian for LSS

The time-dependent translation symmetry laid out above was justified at the level of the equations of motion. It would be useful to see the same at the level of the action. In this section, we provide the action that describes the dark (i.e. pressureless) matter dynamics under gravity.¹⁴ We should stress that, for our discussion of the consistency relation, the action is not strictly necessary; the equations of motion are as good a guide to the symmetry. Moreover, the action we will write down concerns only dark matter; it does not cover realistic observables such as galaxies, while the consistency relation applies regardless of the complex astrophysics that might be present in such observables. Nonetheless, the dark matter action is useful for conceptual understanding. We provide it here for completeness, and connect it with a more well known fluid action in Appendix A. For simplicity, we assume potential flow; an extension to allow for vorticity should be straightforward, along the lines of [30]. Readers not interested in the action perspective can skip to §3 – the rest of the paper does not depend on this section.

Let us motivate the construction of the action by reducing the standard pressureless LSS equations (9) into a single equation for the velocity potential π . The Euler equation can be integrated once to give:

$$\Phi = -\frac{1}{a} \left[(a\pi)' + \frac{1}{2}a(\nabla\pi)^2 \right], \quad (48)$$

where $(\nabla\pi)^2$ stands for $\partial_i\pi\partial_i\pi$.¹⁵ The Poisson equation then gives us:

$$\delta = -\frac{2M_P^2}{\bar{\rho}a^3}\nabla^2 \left[(a\pi)' + \frac{1}{2}a(\nabla\pi)^2 \right]. \quad (49)$$

The continuity equation can thus be turned into a single equation for π :

$$-2M_P^2\nabla^2 \left[(a\pi)' + \frac{1}{2}a(\nabla\pi)^2 \right]' + \bar{\rho}a^3\nabla^2\pi - 2M_P^2\nabla_i \left(\nabla_i\pi\nabla^2 \left[(a\pi)' + \frac{1}{2}a(\nabla\pi)^2 \right] \right) = 0. \quad (50)$$

This is a complicated looking equation, but it is not too difficult to guess the form of the associated action:

$$S = - \int d^4x \left[\frac{1}{2}\bar{\rho}a^4(\nabla\pi)^2 + M_P^2 \left(\nabla \left[(a\pi)' + \frac{1}{2}a(\nabla\pi)^2 \right] \right)^2 \right]. \quad (51)$$

¹³ The soft-mode, by the assumption of growing mode initial condition, is potential flow, but the hard modes need not be, and can even involve multiple streams. See footnote 2.

¹⁴We are indebted to Alberto Nicolis and Filippo Vernizzi who showed us this action.

¹⁵We use ∇_i and ∂_i interchangeably, preferring the former where there is the danger of confusing ∂ with the space-time derivative.

The overall normalization (and sign) is arbitrary from the point of view of reproducing the desired equation of motion, but is chosen to conform to a more general action discussed in Appendix A. It is straightforward to check that this action is invariant under the time-dependent translation symmetry discussed earlier, namely:

$$\eta \rightarrow \tilde{\eta} = \eta \quad , \quad x^i \rightarrow \tilde{x}^i = x^i + n^i \quad , \quad \pi \rightarrow \tilde{\pi} = \pi + n^{i'} x^i . \quad (52)$$

The dynamics of the velocity potential π is completely fixed by this action. From this point of view, the π equation of motion has the interpretation of the continuity equation, if δ is *defined* by Eq. (49); the gravitational potential Φ is *defined* by Eq. (48) so as to reproduce the Poisson equation. With this understanding, the action takes a fairly simple form:

$$S = \int d^4x \frac{\bar{\rho} a^4}{2} (\Phi \delta - \bar{v}^2) , \quad (53)$$

i.e. the Lagrangian is the difference between what resembles potential energy and kinetic energy, though with an unexpected overall sign, which can be understood from the larger context of a fluid with pressure (see Appendix A).

3 Consistency Relations from Diffeomorphisms – General Relativistic Symmetries

The time-dependent translation symmetry noted by KRPP (Eq. 10) appears to be a global symmetry of the Newtonian LSS equations. (Or, more generally, the time-dependent translation as described by Eq. 40 is a symmetry of the equations of motion for dark matter and galaxies.) Our goal in this section is to place it in a larger context: the claim is that this symmetry is actually part of a diffeomorphism in the context of general relativity. This perspective is useful for two reasons: first, it helps us make contact with the earlier work on consistency relations in inflation, which are based on diffeomorphism invariance; second, diffeomorphism invariance allows us to systematically write down further consistency relations. The earlier work generally uses the ζ -gauge, alternatively referred to as the unitary or comoving gauge. On the other hand, in LSS studies, the Newtonian gauge is the more natural one to use. Here, we take advantage of the fact that the full list of consistency relations are already known in the unitary or ζ -gauge [7], and transform each known symmetry in ζ -gauge into a symmetry in the Newtonian gauge. This way, we will obtain an infinite tower of consistency relations in the Newtonian gauge. We emphasize that we could equally well proceed by directly working in the Newtonian gauge, and obtain the same results (see [16] on the dilation and special conformal consistency relations obtained this way). One might wonder why writing down consistency relations in the Newtonian gauge is useful if we already know what they are in the unitary gauge. It has to do with the taking of the Newtonian limit, a subject we will discuss later in this section, and in §4.

In the interest of generality, we allow the presence of multiple components of which pressureless matter/dust is one. We assume adiabatic initial conditions in the sense that all components fluctuate in the same way in the long wavelength limit: in particular their velocity potentials coincide in this limit. We give in §3.1 the general prescription for transforming symmetries known in the unitary gauge to symmetries in Newtonian gauge. In §3.2 we focus on the dilation and the special conformal symmetries,

which are the symmetries that generate only scalar modes, and we show how the KRPP Newtonian consistency relation arises as the sub-Hubble limit of the latter. We comment on the robustness and limitations of the consistency relations in §3.3, adding a relativistic twist to some of the comments made earlier. We also discuss the taking of the Newtonian/sub-Hubble limit. We close with §3.4 on further consistency relations that form an infinite tower – they generally involve the tensor modes. We comment on why there is no useful sub-Hubble limit in these cases.

3.1 Symmetry Transformations from Diffeomorphisms

Here we are interested in symmetry transformations coming from residual gauge/coordinate transformations (i.e. diffeomorphisms)

$$x'^{\mu} = x^{\mu} + \xi^{\mu}, \quad (54)$$

that are allowed even after we have applied the usual gauge-fixing. In the context of inflation, a common gauge is the unitary or ζ -gauge:

$$ds^2 = \dots + a^2 e^{2\zeta} (e^{\gamma})_{ij} dx^i dx^j, \delta\phi = 0, \quad (55)$$

where we have omitted the time-time and time-space components of the metric which are obtainable from the given space-space parts by solving the Hamiltonian and momentum constraints. Here, ζ represents the scalar perturbation and the transverse traceless γ_{ij} represents the tensor perturbation. Vector perturbations are ignored because they are not generated by single field models (a brief discussion of vector modes can be found in Appendix D). The equal time surface is chosen so that the matter field, which we have called ϕ , has no spatial fluctuation. For our application, there can in general be multiple components, in which case $\delta\phi$ is chosen to vanish for one of them. To be concrete, let us choose this to be the dark matter fluid, i.e. we model it as a fluid described by a Lagrangian of the form $P(X)$, where P is some function of $X \equiv -(\partial\phi)^2$. The velocity potential π is related to ϕ by $\delta\phi = \phi - \bar{\phi} = -\bar{\phi}'\pi$, where $\bar{\phi}$ is the background, and $\bar{\phi}'$ is its conformal time derivative (see Appendix A).¹⁶

The full list of residual diffeomorphisms that respect the unitary gauge is worked out in H2K. Since the unitary gauge is a complete gauge-fixing for diffeomorphisms that vanish at spatial infinity, the residual diffeomorphisms must be those that do not vanish at infinity. They take the form:

$$\xi^0 = 0, \quad \xi^i = \xi_{\text{unit.}}^i \sim x^n. \quad (56)$$

No time-diffeomorphism is allowed since that would violate the $\delta\phi = 0$ (or $\pi = 0$) unitary gauge condition, and the allowed spatial diffeomorphism, which we refer to as $\xi_{\text{unit.}}^i$, goes like x^n , where $n = 1, 2, \dots$. We will give explicit expressions for $\xi_{\text{unit.}}^i$ later. They satisfy:¹⁷

$$\text{scalar} + \text{tensor symmetries: } \nabla^2 \xi_{\text{unit.}}^i + \frac{1}{3} \partial_i (\partial_k \xi_{\text{unit.}}^k) = 0. \quad (57)$$

¹⁶By describing the dark matter using a single fluid field ϕ , we are ignoring orbit-crossing and also vorticity. As is clear in the Newtonian discussion, neither one of these assumptions is strictly necessary. We make them only to simplify the general relativistic discussion.

¹⁷This holds only to the lowest order in tensor modes. See H2K, and discussion in footnote 32.

This set of symmetries contains subsets that only generate (nonlinearly) scalar modes, and subsets that only generate tensor modes:

$$\begin{aligned} \text{scalar symmetries : } & \partial_i \xi_{\text{unit.}}^j + \partial_j \xi_{\text{unit.}}^i - \frac{2}{3} \delta_{ij} \partial_k \xi_{\text{unit.}}^k = 0 \\ \text{tensor symmetries : } & \partial_i \xi_{\text{unit.}}^i = 0, \quad \nabla^2 \xi_{\text{unit.}}^i = 0. \end{aligned} \quad (58)$$

The spatial diffeomorphism $\xi_{\text{unit.}}^i$ can be considered to be time-independent.¹⁸

In LSS studies, it is more common to employ the Newtonian gauge instead:

$$ds^2 = a^2 \left[-(1 + 2\Phi) d\eta^2 + 2S_i dx^i d\eta + ((1 - 2\Psi)\delta_{ij} + \gamma_{ij}) dx^i dx^j \right], \quad (59)$$

where we no longer impose $\pi = 0$, Φ and Ψ are the scalar modes, the transverse traceless γ_{ij} denotes the tensor modes as before, and the divergence-free S_i represents the vector modes (which is set to zero in this paper). Here, we work perturbatively in the metric perturbations, since the Newtonian-gauge metric perturbations are expected to be small even in the highly nonlinear regime where the density fluctuation δ is large, and including higher order metric perturbations corrects the consistency relations by negligible amounts.¹⁹ Under a small diffeomorphism ξ^μ , the nonlinear transformations of the metric fluctuations are:²⁰

$$\begin{aligned} \Delta_{\text{nl.}} \Phi &= -\xi^{0'} - \mathcal{H} \xi^0, \quad \Delta_{\text{nl.}} \Psi = \mathcal{H} \xi^0 + \frac{1}{3} \partial_k \xi^k, \\ \Delta_{\text{nl.}} g_{0i} &= a^2 (\partial_i \xi^0 - \partial_0 \xi^i), \quad \Delta_{\text{nl.}} g_{ij} - \frac{1}{3} \delta_{ij} \Delta_{\text{nl.}} g_{kk} = -a^2 \left(\partial_i \xi^j + \partial_j \xi^i - \frac{2}{3} \delta_{ij} \partial_k \xi^k \right). \end{aligned} \quad (60)$$

Given each symmetry in the unitary gauge, it is straightforward to deduce the corresponding symmetry in the Newtonian gauge. Let us break it down into a number of steps. First, we begin with the metric in Newtonian gauge, where $\pi = \pi_0 \neq 0$. We assume $\Psi = \Phi$, in the absence of anisotropic stress.²¹ To convert to the unitary gauge, we apply a time-diffeomorphism $\xi^0 = -\pi_0$ to make the scalar field ϕ spatially homogeneous. Second, we apply the known unitary-gauge symmetry transformation $\xi^i = \xi_{\text{unit.}}^i$. Third, we wish to return to Newtonian gauge. The first and second steps in general make $\Psi \neq \Phi$. To restore equality, we apply an additional time-diffeomorphism $\xi^0 = \pi_0 + \xi_{\text{add.}}^0$. We also need to ensure $g_{0i} = 0$ (no vector modes²²), and thus an additional spatial diffeomorphism $\xi_{\text{add.}}^i$ may be necessary. It is shown in Appendix B that the requisite additional time- and space-diffeomorphisms are:

$$\xi_{\text{add.}}^0 = -\frac{1}{3c} D' \partial_i \xi_{\text{unit.}}^i, \quad \xi_{\text{add.}}^i = \frac{1}{c} D \nabla^2 \xi_{\text{unit.}}^i. \quad (61)$$

¹⁸Adiabatic mode conditions in the unitary gauge actually make $\xi_{\text{unit.}}^i$ time-dependent in general. As shown in H2K, its time-independent part alone is sufficient to deduce the consistency relations. We will implement the adiabatic mode conditions separately in the Newtonian gauge computation.

¹⁹See footnote 32 for a more detailed discussion of this point.

²⁰The net (linear + nonlinear) transformation of the metric is given by $\Delta g_{\mu\nu} = -\xi^\alpha \partial_\alpha g_{\mu\nu} - g_{\alpha\mu} \partial_\nu \xi^\alpha - g_{\alpha\nu} \partial_\mu \xi^\alpha$.

²¹This is an adiabatic mode condition in the Newtonian gauge. See discussions in Appendix B.

²²The absence of vector modes is assumed in two places. Assuming $\nabla^2 \xi_{\text{unit.}}^i + \partial_i (\partial \cdot \xi)/3 = 0$ means there is no vector mode in the spatial part of the metric. In addition, our choice of $\xi_{\text{add.}}^\mu$ ensures there is no vector mode in the space-time part of the metric either.

Here, D is the linear growth factor satisfying the following equation:

$$D'' + 2\mathcal{H}D' - c = 0, \quad (62)$$

where c is a constant (independent of time and space). In other words, the following diffeomorphism is a symmetry of Newtonian gauge:

$$\xi^0 = \xi_{\text{add}}^0, \quad \xi^i = \xi_{\text{unit.}}^i + \xi_{\text{add.}}^i, \quad (63)$$

where $\xi_{\text{unit.}}^i$ is the residual (time-independent) diffeomorphism allowed by the unitary gauge (Eqs. 57 & 58). Furthermore, it can be shown that this diffeomorphism satisfies the adiabatic mode conditions, i.e. the perturbations that are nonlinearly generated match the time-dependence of very soft (growing) physical modes. This is why the linear growth factor D appears in the diffeomorphism. The derivation is given in Appendix B. (The attentive reader might wonder why the linear growth factor D – a quantity that shows up in the Newtonian discussion of sub-Hubble perturbations – appears also in a general relativistic discussion, and how Eq. (62) is related to the more familiar growth equation (Eq. 14). This is discussed in Appendix C). An important underlying assumption is that all fluid components move with the same velocity in the soft limit. Under this assumption, it is shown in Appendix C that the velocity, or velocity potential π , evolves as:

$$\lim_{\vec{q} \rightarrow 0} \pi_{\vec{q}} \propto D'. \quad (64)$$

In the context of a general relativistic discussion, this statement (strictly speaking) holds in the super-Hubble limit $q \ll \mathcal{H}$. What is interesting is that for the $\pi_{\vec{q}}$ of pressureless matter, this statement holds also for sub-Hubble (but linear) scales. It is this fact that makes an interesting Newtonian consistency relation possible.²³

For the purpose of deducing the consistency relations, we also need to know how other LSS observables transform under a diffeomorphism. From the way a scalar should transform, one can see the velocity potential $\pi \equiv -\delta\phi/\bar{\phi}'$ should transform by

$$\Delta\pi = \Delta_{\text{lin.}}\pi + \Delta_{\text{nl.}}\pi \quad \text{where} \quad \Delta_{\text{lin.}}\pi = -\xi^0 \frac{(\bar{\phi}'\pi)'}{\bar{\phi}'} - \xi^i \partial_i \pi, \quad \Delta_{\text{nl.}}\pi = \xi^0. \quad (65)$$

We will mostly need only the nonlinear part of the π transformation. As emphasized above, the assumption of potential flow is not strictly necessary. The nonlinear transformation of the velocity can also be deduced by transforming the 4-velocity U^μ .²⁴

$$\Delta_{\text{nl.}}v^i = \xi^{i'}. \quad (66)$$

Another LSS observable of interest is the mass density fluctuation δ . Its transformation is:

$$\Delta\delta = \Delta_{\text{nl.}}\delta + \Delta_{\text{lin.}}\delta, \quad \Delta_{\text{nl.}}\delta = -\xi^0 \frac{\bar{\rho}'}{\bar{\rho}}, \quad \Delta_{\text{lin.}}\delta = -\xi^\mu \partial_\mu \delta - \xi^0 \frac{\bar{\rho}'}{\bar{\rho}} \delta. \quad (67)$$

²³The fact that the soft π is proportional to D' is nicely consistent with $\xi^0 \propto D'$, since $\Delta_{\text{nl.}}\pi = \xi^0$ (see Eq. 65).

²⁴One can use $U^\mu = (1 - \Phi, v^i)/a$, valid to the lowest order in velocity and perturbations, with the understanding that $v^i = dx^i/d\eta$. In this paper, by relativistic effects, we are generally interested in effects on super-Hubble scales as opposed to effects associated with high peculiar velocities.

One could set $-\bar{\rho}'/\bar{\rho} = 3\mathcal{H}$ for $\bar{\rho}$ that redshifts like pressureless matter, but we will keep the discussion general. The generalization to the galaxy density fluctuation δ_g (or the fluctuation of any component) is immediate:

$$\Delta\delta_g = \Delta_{\text{nl.}}\delta_g + \Delta_{\text{lin.}}\delta_g \quad , \quad \Delta_{\text{nl.}}\delta_g = -\xi^0\frac{\bar{\rho}'_g}{\bar{\rho}_g} \quad , \quad \Delta_{\text{lin.}}\delta_g = -\xi^\mu\partial_\mu\delta_g - \xi^0\frac{\bar{\rho}'_g}{\bar{\rho}_g}\delta_g \quad , \quad (68)$$

where $\bar{\rho}_g$ is the mean galaxy number density. In both cases, the linear part of the transformations would resemble more what one expects for a scalar if we consider $\delta\rho = \bar{\rho}\delta$ instead of δ :

$$\Delta_{\text{nl.}}\delta\rho = -\xi^0\bar{\rho}' \quad , \quad \Delta_{\text{lin.}}\delta\rho = -\xi^\mu\partial_\mu\delta\rho \quad . \quad (69)$$

3.2 Scalar Consistency Relations

Let us first derive the consistency relations that involve only scalar modes, i.e. where only scalar modes are nonlinearly generated. Recall from §3.1 that the scalar symmetries take the form:

$$\xi^0 = \xi_{\text{add.}}^0 \quad , \quad \xi^i = \xi_{\text{unit.}}^i + \xi_{\text{add.}}^i \quad , \quad (70)$$

with $\xi_{\text{unit.}}^i$ and $\xi_{\text{add.}}^\mu$ satisfying:

$$\partial_i\xi_{\text{unit.}}^j + \partial_j\xi_{\text{unit.}}^i - \frac{2}{3}\delta_{ij}\partial_k\xi_{\text{unit.}}^k = 0 \quad , \quad (71)$$

$$\xi_{\text{add.}}^0 = -\frac{1}{3c}D'\partial_i\xi_{\text{unit.}}^i \quad , \quad \xi_{\text{add.}}^i = \frac{1}{c}D\nabla^2\xi_{\text{unit.}}^i \quad , \quad (72)$$

where D is the linear growth factor obeying $D'' + 2\mathcal{H}D' - c = 0$, with c being a constant.

As discussed before, since the unitary gauge is a complete gauge-fixing for diffeomorphisms that vanish at spatial infinity, the residual diffeomorphism of interests must be one where $\xi_{\text{unit.}}^i$ does not vanish at infinity. Following H2K, we can express $\xi_{\text{unit.}}^i$ as a power series:

$$\xi_{\text{unit.}}^i = \sum_{n=0}^{\infty} \frac{1}{(n+1)!} M_{i\ell_0\dots\ell_n} x^{\ell_0} \dots x^{\ell_n} \quad , \quad (73)$$

where each $M_{i\ell_0\dots\ell_n}$ represents a constant coefficient, symmetric in its last $n+1$ indices. As pointed out by [5, 6], the only scalar symmetries are those associated with $n=0$: $\xi_{\text{unit.}}^i \sim x$ (dilation) and $n=1$: $\xi_{\text{unit.}}^i \sim x^2$ (special conformal transformation).

3.2.1 The Dilation Consistency Relation

Dilation is described by $\xi_{\text{unit.}}^i = \lambda x^i$ where λ is a constant. Plugging this into Eq. (72) tells us $\xi_{\text{add.}}^0 = -(\lambda/c)D'$ and $\xi_{\text{add.}}^i = 0$. In other words, the net residual diffeomorphism in Newtonian gauge is

$$\xi^0 = \epsilon \quad , \quad \xi^i = \lambda x^i \quad \text{with} \quad \epsilon \equiv -\frac{\lambda}{c}D' \quad , \quad (74)$$

where λ is a constant. This symmetry involves a spatial dilation + an accompanying time translation, with the two related by a differential equation: $\epsilon' + 2\mathcal{H}\epsilon + \lambda = 0$. We will refer to the resulting consistency

relation simply as the dilation consistency relation, even though the symmetry involves more than spatial dilation. To deduce the associated consistency relation, we employ Eq. (22). Two pieces of information are needed to use it. One is the nonlinear shift of π in Fourier space, obtained by taking the Fourier transform of Eq. (65):

$$\Delta_{\text{nl.}}\pi_{\vec{q}}^* = (2\pi)^3 \delta_D(\vec{q}) \epsilon. \quad (75)$$

The other piece of information we need is the linear transformation of the high momentum observable(s). Here, let us use the density fluctuation $\delta_{\vec{k}}$ as the observable at high momentum. By a Fourier transform of Eq. (67), we find

$$\Delta_{\text{lin.}}\delta_{\vec{k}} = \left[-\epsilon \left(\frac{\bar{\rho}'}{\bar{\rho}} + \partial_{\eta} \right) + \lambda(3 + \vec{k} \cdot \partial_{\vec{k}}) \right] \delta_{\vec{k}}. \quad (76)$$

Plugging these two pieces into the master equation (22), we see that

$$\lim_{\vec{q} \rightarrow 0} \frac{\langle \pi_{\vec{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^c}{P_{\pi}(q)} \epsilon(\eta) = \sum_{a=1}^N \left[-\epsilon(\eta_a) \left(\frac{\bar{\rho}'}{\bar{\rho}} \Big|_{\eta_a} + \partial_{\eta_a} \right) + \lambda \left(3 + \vec{k}_a \cdot \partial_{\vec{k}_a} \right) \right] \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^c, \quad (77)$$

where the time-dependence should be understood as follows: the soft \vec{q} mode is evaluated at time η , while the hard mode \vec{k}_a is evaluated at time η_a , meaning each hard mode can be at a different time. This is why the ϵ on the left is at time η – it is associated with the nonlinear shift in π and therefore the soft mode – and the ϵ 's on the right are evaluated at the respective η_a , since each is associated with the linear transformation of the corresponding hard mode.

The connected N- and (N+1)-point functions on both sides contain the momentum conserving delta function. Its removal requires some care since the derivatives with respect to momentum on the right hand side act on the delta function:

$$\sum_{a=1}^N \vec{k}_a \cdot \partial_{\vec{k}_a} \delta_D(\vec{k}_1 + \dots + \vec{k}_N) = -3\delta_D(\vec{k}_1 + \dots + \vec{k}_N), \quad (78)$$

which can be established by rewriting the delta function as $(2\pi)^{-3} \int d^3x e^{i(\vec{k}_1 + \dots + \vec{k}_N) \cdot \vec{x}}$, and integrating by parts. Thus, removing the delta function on both sides, with $\langle \dots \rangle^{c'}$ representing the connected correlation function without δ_D , we have²⁵

$$\lim_{\vec{q} \rightarrow 0} \frac{\langle \pi_{\vec{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}}{P_{\pi}(q)} \epsilon(\eta) = \left(3\lambda(N-1) + \sum_{a=1}^N \left[-\epsilon(\eta_a) \left(\frac{\bar{\rho}'}{\bar{\rho}} \Big|_{\eta_a} + \partial_{\eta_a} \right) + \lambda \vec{k}_a \cdot \partial_{\vec{k}_a} \right] \right) \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}, \quad (79)$$

with the understanding that ϵ and (the constant) λ are related by Eq. (74), and where the N-point function depends on the time associated with each of the N modes. Using the relation, we can rewrite the *dilation consistency relation* as

$$\boxed{\lim_{\vec{q} \rightarrow 0} \frac{\langle \pi_{\vec{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}}{P_{\pi}(q)} = - \left(\frac{3c}{D'(\eta)}(N-1) + \sum_{a=1}^N \left[\frac{D'(\eta_a)}{D'(\eta)} \left(\frac{\bar{\rho}'}{\bar{\rho}} \Big|_{\eta_a} + \partial_{\eta_a} \right) + \frac{c}{D'(\eta)} \vec{k}_a \cdot \partial_{\vec{k}_a} \right] \right) \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'},} \quad (80)$$

²⁵ Note that $\langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}$ should be understood to be a function of only $N-1$ momenta. For instance, we can think of $\vec{k}_N = -\vec{k}_1 - \vec{k}_2 \dots - \vec{k}_{N-1}$. Thus the derivative $\partial_{\vec{k}_N}$, keeping $\vec{k}_1, \dots, \vec{k}_{N-1}$ fixed, vanishes. This point is particularly important for the higher consistency relations. See [11].

with the understanding that $c = D'' + 2\mathcal{H}D'$ is a constant. It is trivial to generalize the consistency relation by changing the hard modes from δ for the mass density to δ_g for the galaxy density: simply change the mean mass density $\bar{\rho}$ on the right hand side to the mean galaxy number density $\bar{\rho}_g$. This consistency relation can be further rewritten in different forms. We will postpone this discussion until after we discuss the special conformal consistency relation.

3.2.2 The Special Conformal Consistency Relation – Containing the Newtonian Translation Consistency Relation

Next, we consider the special conformal transformation: $\xi_{\text{unit.}}^i = 2\vec{b} \cdot \vec{x} x^i - b^i \vec{x} \cdot \vec{x}$, where \vec{b} is a constant vector. Plugging this into Eq. (72), we see that the requisite accompanying diffeomorphism is: $\xi_{\text{add.}}^0 = -(2/c)D'\vec{b} \cdot \vec{x}$ and $\xi_{\text{add.}}^i = -(2/c)Db^i$. Putting everything together, we see that the symmetry is:

$$\xi^0 = n^{i'} x^i, \quad \xi^i = n^i + 2\vec{b} \cdot \vec{x} x^i - b^i \vec{x} \cdot \vec{x} \quad \text{with} \quad n^i \equiv -\frac{2}{c}Db^i \quad (81)$$

where \vec{b} is a constant vector. We refer to the implied consistency relation as the special conformal consistency relation, even though the full symmetry transformation involves a time diffeomorphism and a spatial translation in addition to the special conformal transformation – these transformations are related via $n^{i''} + 2\mathcal{H}n^{i'} + 2b^i = 0$. As we will see, the Newtonian translation consistency relation is contained in here.

Once again, we employ the master equation (Eq. 22), for which we need the nonlinear transformation of the velocity potential π and the linear transformation of the hard modes – as in the case of dilation, we choose the observable to be the density fluctuation δ for the hard modes. Under the current symmetry transformation, we have:

$$\Delta_{\text{nl.}}\pi_{\vec{q}}^* = i\vec{n}' \cdot \partial_{\vec{q}}[(2\pi)^3 \delta_D(\vec{q})], \quad (82)$$

$$\Delta_{\text{lin.}}\delta_{\vec{k}} = i \left[\vec{n}' \cdot \partial_{\vec{k}} \left(\frac{\bar{\rho}'}{\bar{\rho}} + \partial_{\eta} \right) + \vec{n} \cdot \vec{k} - (6\vec{b} \cdot \partial_{\vec{k}} + 2b^j k^i \partial_{k^j} \partial_{k^i} - \vec{b} \cdot \vec{k} \nabla_{\vec{k}}^2) \right] \delta_{\vec{k}}. \quad (83)$$

Substituting the above into Eq. (22), we see that the left hand side (LHS) is

$$\begin{aligned} \text{LHS} = & \lim_{\vec{q} \rightarrow 0} (-i\vec{n}'(\eta)) \cdot \partial_{\vec{q}} \left[(2\pi)^3 \delta_D(\vec{q} + \vec{k}_1 + \dots + \vec{k}_N) \right] \frac{\langle \pi_{\vec{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}}{P_{\pi}(q)} \\ & - i(2\pi)^3 \delta_D(\vec{q} + \vec{k}_1 + \dots + \vec{k}_N) \vec{n}'(\eta) \cdot \partial_{\vec{q}} \left[\frac{\langle \pi_{\vec{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}}{P_{\pi}(q)} \right]. \end{aligned} \quad (84)$$

The right hand side (RHS) is

$$\text{RHS} = \sum_{a=1}^N i \left[\vec{n}'(\eta_a) \cdot \partial_{\vec{k}_a} \left(\frac{\bar{\rho}'}{\bar{\rho}} \Big|_{\eta_a} + \partial_{\eta_a} \right) + \vec{n}(\eta_a) \cdot \vec{k}_a - (6\vec{b} \cdot \partial_{\vec{k}_a} + 2b^i k_a^j \partial_{k_a^j} \partial_{k_a^i} - \vec{b} \cdot \vec{k}_a \nabla_{\vec{k}_a}^2) \right] \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^c. \quad (85)$$

The connected N-point function $\langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^c$ contains an overall momentum conserving delta function. The momentum derivative acts non-trivially on it. To simplify, it is useful to know:

$$\sum_{a=1}^N \left(2b^j k_a^i \partial_{k_a^j} \partial_{k_a^i} - \vec{b} \cdot \vec{k}_a \nabla_{k_a}^2 \right) \delta_D(\vec{k}_{\text{tot.}}) = -6\vec{b} \cdot \partial_{\vec{k}_{\text{tot.}}} \delta_D(\vec{k}_{\text{tot.}}), \quad (86)$$

where $\vec{k}_{\text{tot}} \equiv \vec{k}_1 + \dots + \vec{k}_N$, and which can be proved by rewriting the delta function as the spatial integral of a plane wave. The term $\sum_a 6\vec{b} \cdot \partial_{\vec{k}_a} \delta_D(\vec{k}_{\text{tot.}})$ can be rewritten as $6N\vec{b} \cdot \partial_{\vec{k}_{\text{tot.}}} \delta_D(\vec{k}_{\text{tot.}})$; indeed, any $\partial_{k_a^i} \delta_D(\vec{k}_{\text{tot.}})$ can be written as $\partial_{k_{\text{tot.}}^i} \delta_D(\vec{k}_{\text{tot.}})$. Then there are terms that involve one momentum derivative on the delta function and one momentum derivative on the N-point function:

$$\begin{aligned} & \sum_{a=1}^N 2b^i k_a^j \partial_{k_a^j} \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'} \partial_{k_a^i} \delta_D(\vec{k}_{\text{tot.}}) + 2b^i \partial_{k_a^j} \delta_D(\vec{k}_{\text{tot.}}) \left(k_a^j \partial_{k_a^i} \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'} - k_a^i \partial_{k_a^j} \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'} \right) \\ &= \left[2\vec{b} \cdot \partial_{\vec{k}_{\text{tot.}}} \delta_D(\vec{k}_{\text{tot.}}) \right] \sum_{a=1}^N \vec{k}_a \cdot \partial_{\vec{k}_a} \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}, \end{aligned} \quad (87)$$

where we have used rotational invariance of the N-point function to remove the last two terms on the first line. With this understanding, Eq. (85) can be expressed as

$$\begin{aligned} \text{RHS} &= -i \left[\partial_{k_{\text{tot.}}^i} (2\pi)^3 \delta_D(\vec{k}_{\text{tot.}}) \right] \left[6b^i (N-1) + \sum_{a=1}^N \left(-n^{i'}(\eta_a) \left(\frac{\bar{\rho}'}{\bar{\rho}} \Big|_{\eta_a} + \partial_{\eta_a} \right) + 2b^i \vec{k}_a \cdot \partial_{\vec{k}_a} \right) \right] \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'} \\ &\quad -i(2\pi)^3 \delta_D(\vec{k}_{\text{tot.}}) \sum_{a=1}^N \left[- \left(\frac{\bar{\rho}'}{\bar{\rho}} \Big|_{\eta_a} + \partial_{\eta_a} \right) \vec{n}'(\eta_a) \cdot \partial_{\vec{k}_a} - \vec{n}(\eta_a) \cdot \vec{k}_a \right. \\ &\quad \left. + (6\vec{b} \cdot \partial_{\vec{k}_a} + 2b^i k_a^j \partial_{k_a^j} \partial_{k_a^i} - \vec{b} \cdot \vec{k}_a \nabla_{k_a}^2) \right] \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}, \end{aligned} \quad (88)$$

The first line of the above can be equated with the first line of Eq. (84), since what multiplies the derivative of the delta function on both sides replicates the dilation consistency relation Eq. (79). Note that n^i and b^i are related by Eq. (81). Using this, eliminating the dilation consistency relation from both sides,²⁶ and removing the delta function, we obtain the *special conformal consistency relation*:

$$\boxed{\begin{aligned} \lim_{\bar{q} \rightarrow 0} \partial_{q^i} \left[\frac{\langle \pi_{\bar{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}}{P_\pi(q)} \right] &= - \sum_{a=1}^N \left[\left(\frac{\bar{\rho}'}{\bar{\rho}} \Big|_{\eta_a} + \partial_{\eta_a} \right) \frac{D'(\eta_a)}{D'(\eta)} \partial_{k_a^i} + \frac{D(\eta_a)}{D'(\eta)} k_a^i \right. \\ &\quad \left. + \frac{c}{D'(\eta)} \left(3\partial_{k_a^i} + k_a^j \partial_{k_a^j} \partial_{k_a^i} - \frac{1}{2} k_a^i \nabla_{k_a}^2 \right) \right] \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}, \end{aligned}} \quad (89)$$

with the understanding that $c = D'' + 2\mathcal{H}D'$ is a constant (Eq. 165). Just as in the case of the dilation consistency relation, this consistency relation can be easily generalized to the hard modes being the

²⁶This is a general pattern: one can obtain the correct consistency relation for a given symmetry by using the master equation (22), *and* simply ignoring the delta functions on both sides. At first sight, this might appear dangerous as there are derivatives acting on the delta functions, but they invariably multiply consistency relations from symmetries at the lower levels, and so can be removed.

galaxy overdensity – changing δ to δ_g , and changing $\bar{\rho}$ to $\bar{\rho}_g$. Examining the terms on the right hand side, we see that in the sub-Hubble limit, i.e. $k \gg \mathcal{H}$, the term that dominates on the right hand side is $-\sum_a (D(\eta_a)/D'(\eta)) k_a^i \langle \delta_{\vec{k}_1}^- \dots \delta_{\vec{k}_N}^- \rangle^{c'}$, reproducing the translation consistency relation (Eq. 30) derived from the Newtonian equations. We will have more to say about the non-relativistic limit in §3.3.

At the level of the spatial diffeomorphisms, dilation and special conformal transformations exhaust the list of purely scalar symmetries, since it is only dilation and special conformal transformations that respect the second expression of Eq. (71) and do not generate vector or tensor modes. It is also worth noting that the special conformal transformation consistency relation strictly speaking receives (small) corrections on the right hand side, a point to which we will return (footnote 32).

3.3 Robustness and Limitations of the Consistency Relations - a Relativistic Perspective and the Newtonian Limit

It is useful to pause, and reflect on the fully relativistic consistency relations derived so far. Some of our discussions here mirror the earlier ones in the Newtonian context (§2.2), but with a relativistic twist. We also discuss the issue of taking the Newtonian, i.e. sub-Hubble, limit.

1. Newtonian limit. The special conformal consistency relation Eq. (89) is the relativistic analog of the Newtonian translation consistency relation Eq. (30). The former reduces to the latter in the sense that:

$$\lim_{\vec{q} \rightarrow 0} \frac{\partial}{\partial q_j} \left[\frac{\langle \pi_{\vec{q}} \delta_{\vec{k}_1}^- \dots \delta_{\vec{k}_N}^- \rangle^{c'}}{P_\pi(q)} \right] = - \sum_{a=1}^N \frac{D(\eta_a)}{D'(\eta)} k_{aj} \langle \delta_{\vec{k}_1}^- \dots \delta_{\vec{k}_N}^- \rangle^{c'} \times (1 + O(\mathcal{H}^2/k^2)) , \quad (90)$$

where the \mathcal{H}^2/k^2 -suppressed terms can be ignored in the sub-Hubble limit. Note that the unsuppressed (Newtonian) terms are of the order of $\frac{k}{\mathcal{H}} \langle \delta_{\vec{k}_1}^- \dots \delta_{\vec{k}_N}^- \rangle^{c'}$. Similarly, we can think of the dilation consistency relation Eq. (80) as the relativistic analog of Eq. (28). The dilation consistency relation takes the form:

$$\lim_{\vec{q} \rightarrow 0} \frac{\langle \pi_{\vec{q}} \delta_{\vec{k}_1}^- \dots \delta_{\vec{k}_N}^- \rangle^{c'}}{P_\pi(q)} = O\left(\frac{k}{\mathcal{H}} \langle \delta_{\vec{k}_1}^- \dots \delta_{\vec{k}_N}^- \rangle^{c'}\right) \times O(\mathcal{H}^2/k) = O\left(q \frac{k}{\mathcal{H}} \langle \delta_{\vec{k}_1}^- \dots \delta_{\vec{k}_N}^- \rangle^{c'}\right) \times O(\mathcal{H}/k)(\mathcal{H}/q) , \quad (91)$$

which can be compared against $q \times$ Eq. (90). We can see that the right hand side of the above expression is $O(\mathcal{H}/k)O(\mathcal{H}/q)$ times $q \times$ Eq. (90). In the sub-Hubble limit where \mathcal{H} is small compared to both q and k , it is therefore consistent to think of Eq. (91) as vanishing – reducing to Eq. (28).²⁷

2. Combining consistency relations. It is worth pointing out that, just as in the Newtonian case where Eqs. (28) and (30) can be combined into a single equation (31), the general relativistic dilation

²⁷Eq. (28) was derived using the shift symmetry $\pi \rightarrow \pi + b$, where b is a constant. The reader might wonder how that argument breaks down in the relativistic context. The point is that a constant shift in π has to be accompanied by a time-dependent shift in Φ (see e.g. Eq. 48). Such a time-dependent shift is not a symmetry of the kinetic term for the metric once time-derivatives are taken into account, unless coordinates change too. It is interesting to note that $\pi \rightarrow \pi + b/\bar{\phi}'$ is a symmetry (see Appendix A) because of the shift symmetry in ϕ ; however, this symmetry does not correspond to the growing mode vacuum and therefore does not lead to a consistency relation. See Appendix D for a further discussion.

and special conformal consistency relations can be combined into:

$$\lim_{\vec{q} \rightarrow 0} \frac{\langle \pi_{\vec{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}}{P_\pi(q)} = - \left(\frac{3c}{D'(\eta)} (N-1) + \sum_{a=1}^N \left[\frac{D'(\eta_a)}{D'(\eta)} \left(\frac{\vec{\rho}'}{\bar{\rho}} \Big|_{\eta_a} + \partial_{\eta_a} \right) (1 + \vec{q} \cdot \partial_{\vec{k}_a}) + \frac{D(\eta_a)}{D'(\eta)} \vec{q} \cdot \vec{k}_a \right. \right. \\ \left. \left. + \frac{c}{D'(\eta)} \left(\vec{k}_a \cdot \partial_{\vec{k}_a} + 3\vec{q} \cdot \partial_{\vec{k}_a} + q^i k_a^j \partial_{k_a^j} \partial_{k_a^i} - \frac{1}{2} \vec{q} \cdot \vec{k}_a \nabla_{k_a}^2 \right) \right] \right) \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}, \quad (92)$$

where the constant $c = D'' + 2\mathcal{H}D'$.

3. Alternative pions. Recall that π , δ and Φ all shift nonlinearly under the symmetries of interest.²⁸ One might wonder whether we could have derived the consistency relation with δ or Φ playing the role of the pion instead. The answer is affirmative. Let us compare these nonlinear shifts: $\Delta_{\text{nl.}}\pi = \xi^0$, $\Delta_{\text{nl.}}\delta = -\xi^0 \vec{\rho}'/\bar{\rho}$, $\Delta_{\text{nl.}}\Phi = -\xi^{0'} - \mathcal{H}\xi^0$. Recalling that $\xi^0 \propto D'$, we see that $\Delta_{\text{nl.}}\delta = -\vec{\rho}'/\bar{\rho} \times \Delta_{\text{nl.}}\pi$, and $\Delta_{\text{nl.}}\Phi = -(D'' + \mathcal{H}D')/D' \times \Delta_{\text{nl.}}\pi$. One can thus run the same arguments as before, and arrive at essentially the same consistency relation Eq. (92), with the right hand side unaltered, but the left hand side replaced by

$$\text{LHS} \rightarrow -\frac{\vec{\rho}'}{\bar{\rho}} \Big|_{\eta} \times \lim_{\vec{q} \rightarrow 0} \frac{\langle \delta_{\vec{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}}{P_\delta(q)}, \quad (93)$$

or

$$\text{LHS} \rightarrow -\frac{D'' + \mathcal{H}D'}{D'} \Big|_{\eta} \times \lim_{\vec{q} \rightarrow 0} \frac{\langle \Phi_{\vec{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}}{P_\Phi(q)}. \quad (94)$$

The consistency relations expressed using π , δ or Φ as the soft pion are all equivalent – with one important caveat, which is related to the squeezing constraint.

4. Squeezing constraint. The reader might want to consult §2.2 for a parallel discussion of the squeezing constraint in the Newtonian context. The consistency relation, whether expressed in terms of π as in Eq. (92), or expressed in terms of δ or Φ as in Eq. (93) or (94), is a statement about the $\vec{q} \rightarrow 0$ limit. In the relativistic context, this means, in addition to $q \ll k_1, \dots, k_N$, the soft mode should strictly speaking be super-Hubble, i.e. $q < \mathcal{H}$. On the other hand, in LSS we are typically interested in sub-Hubble modes, so the question arises: under what conditions does the consistency relation remain valid when all modes, including the soft one, are sub-Hubble (while maintaining the hierarchy $q \ll k_1, \dots, k_N$)?

We show in Appendix C a special fact about the velocity potential π for pressureless matter: it has the same time dependence $\propto D'$ (Eqs. 164 & 165) regardless of whether the wave-mode of interest is inside or outside the Hubble radius. (When it is outside the Hubble radius, this statement is true throughout the entire history of the universe; when it is inside the Hubble radius, this statement is strictly true only after radiation domination.) The consistency relation written in terms of the matter π (92) can therefore be safely taken inside the Hubble radius, even for the soft mode, since the same diffeomorphism is capable of generating the correct π regardless of whether it is on super- and sub-Hubble scales. In this limit, we

²⁸We could also discuss the nonlinear shift of δ_n , which coincides with δ for pressureless matter. (See Appendix C.)

recover the Newtonian consistency relation: the term $-\sum_a [D(\eta_a)/D'(\eta)] \vec{q} \cdot \vec{k}_a \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}$ dominates on the right hand side.²⁹ Similar statements hold for Φ as the soft pion.

The same is not true for δ (here, we focus on the *matter* δ as the soft mode): from the continuity equation (143), it is evident that the time dependence of δ (which is the same as δ_n for pressureless matter) depends on whether the wave-mode is inside or outside the Hubble radius. The consistency relation written using δ as the soft mode takes the form of Eq. (93) only for $q < \mathcal{H}$. If the soft δ mode is within the horizon, the continuity equation tells us $\delta_{\vec{q}} = q^2 (D/D') \pi_{\vec{q}}$, and so the left hand side of the consistency relation should read:

$$\text{LHS} \rightarrow \lim_{\vec{q} \rightarrow 0} q^2 \frac{D(\eta)}{D'(\eta)} \times \frac{\langle \delta_{\vec{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}}{P_\delta(q)}, \quad (95)$$

while as discussed above, the right hand side reduces to $-\sum_a [D(\eta_a)/D'(\eta)] \vec{q} \cdot \vec{k}_a \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}$. This reproduces the Newtonian translation consistency relation written in terms of $\delta_{\vec{q}}$ (Eq. 33). To conclude: the consistency relation expressed in terms of a soft $\delta_{\vec{q}}$ takes a different form outside versus inside the Hubble radius i.e. Eq. (93) versus Eq. (95). The consistency relation expressed using the matter π or Φ as the soft pion maintains the same form regardless.³⁰

5. The existence of an interesting Newtonian limit. From the discussion above, we see that the special conformal consistency relation has a non-trivial Newtonian limit (i.e. the right hand side is non-vanishing), whereas the dilation one does not. What is the underlying reason? From Eq. (72), we see that for a given unitary-gauge transformation $\xi_{\text{unit.}}^i$, the corresponding residual diffeomorphism in Newtonian gauge is

$$\xi^0 \sim \mathcal{H}^{-1} \partial_i \xi_{\text{unit.}}^i, \quad \xi^i \sim \xi_{\text{unit.}}^i + \mathcal{H}^{-2} \nabla^2 \xi_{\text{unit.}}^i. \quad (96)$$

The associated consistency relation, making use of the relation $\delta_{\vec{q}} \sim q^2 \pi_{\vec{q}} / \mathcal{H}$ in the sub-Hubble limit, can be written schematically as:

$$\lim_{\vec{q} \rightarrow 0} \frac{q^2}{\mathcal{H}^2} \frac{\langle \delta_{\vec{q}} \delta_{\vec{k}} \dots \rangle^{c'}}{P_\delta(q)} [\partial_i \xi_{\text{unit.}}^i]_q \sim ([\partial_i \xi_{\text{unit.}}^i]_k + k_i [\xi_{\text{unit.}}^i]_k + k_i \mathcal{H}^{-2} [\nabla^2 \xi_{\text{unit.}}^i]_k) \langle \delta_{\vec{k}} \dots \rangle^{c'}. \quad (97)$$

Here, $[\]_k$ denotes the Fourier transform of the quantity of interest at momentum k , with the delta function removed. For instance, for $\xi_{\text{unit.}}^i \sim x^{n+1}$, we have $[\xi_{\text{unit.}}^i]_k \sim k^{-n-1}$, $[\partial_i \xi_{\text{unit.}}^i]_k \sim k^{-n}$, $[\partial_i \xi_{\text{unit.}}^i]_q \sim q^{-n}$, and $[\nabla^2 \xi_{\text{unit.}}^i]_k \sim k^{-n+1}$. This gives

$$\lim_{\vec{q} \rightarrow 0} \frac{\langle \delta_{\vec{q}} \delta_{\vec{k}} \dots \rangle^{c'}}{P_\delta(q)} \sim \left(\frac{q}{k}\right)^n \left(\frac{\mathcal{H}^2}{q^2} + \frac{k^2}{q^2}\right) \langle \delta_{\vec{k}} \dots \rangle^{c'}. \quad (98)$$

²⁹There is one subtlety though: for a wave-mode that enters the Hubble radius during radiation domination, its time-evolution deviates from D' during part of its history, and so strictly speaking the consistency relation does not apply if the soft mode belongs to this category. An alternative way to put it is this: when the wave-mode is within the Hubble radius (or more precisely, within the sound horizon) during radiation domination, neither the matter nor the radiation moves with a velocity that agrees with D' . A diffeomorphism that obeys the adiabatic mode conditions (e.g. Eq. 74 for dilation, or Eq. 81 for special conformal transformation) cannot generate the correct velocity for either component. Even in this case, we expect the consistency relation to still be a good approximation in the late universe, to the extent that most of the late-time non-Gaussianity is generated after radiation domination. We thank Paolo Creminelli for discussions on this point.

³⁰Using the baryon π as the soft pion is permissible too, as long as one stays above the Jeans scale.

In the sub-Hubble (and squeezed) limit where $\mathcal{H} \ll q \ll k$, this suggests we have the dimensionless ratio $\langle \delta_{\vec{q}} \delta_{\vec{k}} \dots \rangle^{c'} / (P_\delta(q) \langle \delta_{\vec{k}} \dots \rangle^{c'}) \sim (q/k)^n (k/q)^2$. For $n = 1$, the special conformal case, this reproduces correctly the Newtonian translation consistency relation. For $n = 0$, the dilation case, this does not work. The reason is that the k^2/q^2 term in Eq. (98), which is the dominant term in the sub-Hubble limit, originates from $\nabla^2 \xi_{\text{unit.}}^i$, which vanishes for the dilation $\xi_{\text{unit.}}^i = \lambda x^i$. Our naïve power-counting argument also suggests there could be additional $n > 1$ consistency relations that are non-trivial in the Newtonian limit. As we will see in the next section, the $n > 1$ consistency relations generally involve tensors, which complicates taking the squeezed mode to within the Hubble radius.

Let us close this section by emphasizing the robustness of the consistency relations. As in the Newtonian derivation, the general relativistic derivation makes no assumptions about the dynamics of the hard modes – all we need to know is how they transform under diffeomorphisms. Thus, we expect the consistency relations to hold even for nonlinear, or astrophysically messy, hard modes (though the right hand side of the consistency relations might need to be modified depending on exactly how the modes of interest transform; see comments after Eq. 80 and in footnote 32). Besides the existence of symmetries, which according to the general relativistic perspective are nothing but residual diffeomorphisms, the two key assumptions are the same as in the Newtonian derivation: single field initial condition and adiabatic mode conditions, in particular that all species move with the same velocity in the soft limit.

3.4 Consistency Relations Involving Tensor Modes

In this section, we move beyond dilation and special conformal transformation to discuss residual diffeomorphisms that generate tensor modes (with or without accompanying scalar modes). We apply the same strategy as the one used for the pure scalar symmetries: use the full set of symmetries derived in the unitary gauge by H2K, and map each to a symmetry in the Newtonian gauge by Eq. (63).

The unitary gauge residual diffeomorphisms can be written as (Eq. 73):

$$\xi_{\text{unit.}}^i = \sum_{n=0}^{\infty} \frac{1}{(n+1)!} M_{i\ell_0 \dots \ell_n} x^{\ell_0} \dots x^{\ell_n}, \quad (99)$$

where the constant coefficients M satisfy:

$$M_{i\ell\ell\ell_2 \dots \ell_n} = -\frac{1}{3} M_{\ell i\ell\ell_2 \dots \ell_n}. \quad (100)$$

This condition is derived by substituting the power series into Eq. (57): $\nabla^2 \xi_{\text{unit.}}^i + \partial_i (\partial_k \xi_{\text{unit.}}^k) / 3 = 0$. Note that M by definition is symmetric in its last $n+1$ indices. Since we are interested in M that generates tensor modes (in addition to possibly scalar modes), we should impose an additional *adiabatic transversality condition*:

$$\hat{q}^i \left[M_{ij \dots}(\hat{q}) + M_{ji \dots}(\hat{q}) - \frac{2}{3} M_{kk \dots}(\hat{q}) \delta_{ij} \right] = 0. \quad (101)$$

This condition can be understood as enforcing that the tensor generated by our diffeomorphism be extensible to the $\vec{q} \rightarrow 0$ limit of a transverse physical tensor mode. Imagine an M that is nearly constant but tapers off to zero at sufficiently large x : while a constant M yields (derivatives of) a delta function

peaked at $\vec{q} = 0$ in Fourier space, a tapering M yields a smoothed-out version thereof. A tensor mode at a small but finite momentum should be transverse to its own momentum. We demand that even as we take the $\vec{q} = 0$ limit (allowing the tapering of M to occur at larger and larger distances), transversality continues to hold, keeping the direction \hat{q} fixed. This is the content of Eq. (101). The choice of \hat{q} is arbitrary; one could for instance choose it to point in the z direction. In addition, if the diffeomorphism generates *only* tensor modes, further conditions on M come from Eq. (58): $\partial_i \xi_{\text{unit.}}^i = 0$ & $\nabla^2 \xi_{\text{unit.}}^i = 0$, which implies

$$M_{\ell\ell_1\ell_2\ldots\ell_n} = 0 \quad \& \quad M_{i\ell\ell_2\ldots\ell_n} = 0, \quad (102)$$

i.e. M is traceless over any pairs of indices, which also trivially satisfies Eq. (100).

As discussed in §3.1, for each unitary gauge residual diffeomorphism, there is a corresponding one in Newtonian gauge:

$$\xi^\mu = \xi_{\text{unit.}}^\mu + \xi_{\text{add.}}^\mu \quad \text{where} \quad \xi_{\text{add.}}^0 = -\frac{1}{3c} D' \partial_i \xi_{\text{unit.}}^i, \quad \xi_{\text{add.}}^i = \frac{1}{c} D \nabla^2 \xi_{\text{unit.}}^i, \quad (103)$$

with D being the linear growth factor and where c is a constant satisfying $D'' + 2\mathcal{H}D' - c = 0$. Thus, at the level n , the Newtonian gauge diffeomorphism is:

$$\xi^0 = -\frac{D'}{3cn!} M_{\ell\ell_1\ldots\ell_n} x^{\ell_1} \ldots x^{\ell_n}, \quad \xi^i = \frac{1}{(n+1)!} M_{i\ell_0\ell_1\ldots\ell_n} x^{\ell_0} \ldots x^{\ell_n} + \frac{D}{c(n-1)!} M_{i\ell\ell_2\ldots\ell_n} x^{\ell_2} \ldots x^{\ell_n}. \quad (104)$$

This expression holds for all n , with the exception of $n = 0$, in which case the last term on the right for ξ^i is absent. One thing which is immediately clear is that for purely tensor symmetries — diffeomorphisms that generate only tensor modes — we have $\xi_{\text{add.}}^0 = \xi_{\text{add.}}^i = 0$, and so they are identical in the unitary gauge and the Newtonian gauge, as expected. It is worth emphasizing that Eqs. (104) and (100) are general: they apply to symmetries that generate only scalar modes, or only tensor modes, or both. For symmetries that generate tensor modes, adiabatic transversality expressed in Eq. (101) is an additional requirement, and Eq. (102) applies if *only* tensor modes are generated.

To derive the corresponding consistency relations, we need a master equation analogous to Eq. (22) but generalized to allow for the possibility of tensor modes:

$$\int \frac{d^3q}{(2\pi)^3} \left[\frac{\langle \pi_{\vec{q}} \mathcal{O}_{\vec{k}_1} \ldots \mathcal{O}_{\vec{k}_N} \rangle^c}{P_\pi(q)} \Delta_{\text{nl.}} \pi_{\vec{q}}^* + \sum_s \frac{\langle \gamma_{\vec{q}}^s \mathcal{O}_{\vec{k}_1} \ldots \mathcal{O}_{\vec{k}_N} \rangle^c}{P_\gamma(q)} \Delta_{\text{nl.}} \gamma_{\vec{q}}^{s*} \right] = \Delta_{\text{lin.}} \langle \mathcal{O}_{\vec{k}_1} \ldots \mathcal{O}_{\vec{k}_N} \rangle^c. \quad (105)$$

Here, the label s denotes one of the two possible tensor polarization states; a given tensor perturbation $\gamma_{ij}(\vec{q})$ can be decomposed as $\gamma_{ij}(\vec{q}) = \sum_s \epsilon_{ij}^s(\hat{q}) \gamma_{\vec{q}}^s$, where the symmetric traceless polarization tensor $\epsilon_{ij}^s(\hat{q})$ obeys $\hat{q}^i \epsilon_{ij}^s(\hat{q}) = 0$ and $\epsilon_{ij}^s(\hat{q}) \epsilon_{ij}^{s'}(\hat{q})^* = 2\delta^{ss'}$. The tensor power spectrum is defined by $\langle \gamma_{\vec{q}}^s \gamma_{\vec{q}'}^{s'} \rangle = (2\pi)^3 \delta_D(\vec{q} + \vec{q}') \delta^{ss'} P_\gamma(q)$. Note that

$$\Delta_{\text{nl.}} \gamma_{ij} = - \left(\partial_i \xi^j + \partial_j \xi^i - \frac{2}{3} \delta_{ij} \partial_k \xi^k \right). \quad (106)$$

Eq. (105) can alternatively be written as:

$$\int \frac{d^3q}{(2\pi)^3} \left[\frac{\langle \pi_{\vec{q}} \mathcal{O}_{\vec{k}_1} \ldots \mathcal{O}_{\vec{k}_N} \rangle^c}{P_\pi(q)} \Delta_{\text{nl.}} \pi_{\vec{q}}^* + \frac{1}{2} \frac{\langle \gamma_{ij}(\vec{q}) \mathcal{O}_{\vec{k}_1} \ldots \mathcal{O}_{\vec{k}_N} \rangle^c}{P_\gamma(q)} \Delta_{\text{nl.}} \gamma_{ij}(\vec{q})^* \right] = \Delta_{\text{lin.}} \langle \mathcal{O}_{\vec{k}_1} \ldots \mathcal{O}_{\vec{k}_N} \rangle^c, \quad (107)$$

using the fact that $\Delta_{\text{nl.}\gamma_{\vec{q}}}^{s*} = \Delta_{\text{nl.}\gamma_{ij}}(\vec{q})^* \epsilon_{ij}^s(\hat{q})/2$. Let us step through a few low n examples to get a feel for the kind of consistency relations that arise from these diffeomorphisms. The discussion follows that of H2K, with suitable deformations to the Newtonian gauge.

For $n = 0$, $M_{i\ell_0}$ can be written as a trace (dilation), an antisymmetric part (which does not generate a nonlinear shift in the metric³¹) and a symmetric traceless part (anisotropic rescaling, which generates tensor perturbations). Note that $n = 0$ is a special case, in the sense that the last term of Eq. (104) does not exist (because $\nabla^2 \xi_{\text{unit.}}^i = 0$). Focusing on a symmetric traceless $M_{i\ell_0}$, there are 5 independent components. Imposing the adiabatic transversality condition $\hat{q}^i M_{i\ell_0}(\hat{q}) = 0$ reduces the number of independent tensor modes to 2. Thus, at the level of $n = 0$, we have one pure scalar and 2 pure tensor symmetries. It is straightforward to infer the corresponding symmetries in Newtonian gauge: dilation gets deformed as discussed in §3.2.1; the purely tensor symmetries take exactly the same form in the two gauges. The $n = 0$ anisotropic rescaling tensor consistency relation reads:

$$\lim_{\vec{q} \rightarrow 0} \frac{\langle \gamma_{\vec{q}}^s \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}}{P_\gamma(q)} = -\frac{1}{2} \epsilon_{ij}^s(\hat{q})^* \sum_{a=1}^N k_a^i \partial_{k_a^j} \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}, \quad (108)$$

a relation first pointed out by Maldacena [3]. To derive this, we use $\Delta_{\text{nl.}\gamma_{i\ell_0}}(\vec{q})^* = -2M_{i\ell_0}(2\pi)^3 \delta_D(\vec{q})$, which gives $\Delta_{\text{nl.}\gamma_{\vec{q}}}^{s*} = -2(2\pi)^3 \delta_D(\vec{q})$ if we choose $M_{i\ell_0} = \epsilon_{i\ell_0}^s(\hat{q})^*$; the linear transformation of δ under the diffeomorphism $\xi^i = M_{i\ell_0} x^{\ell_0}$ is $\Delta_{\text{lin.}} \delta_{\vec{k}} = M_{i\ell_0} k^i \partial_{k^{\ell_0}} \delta_{\vec{k}}$.

For $n = 1$, there are 3 purely scalar symmetries (the special conformation transformations) and 4 purely tensor ones. The special conformal transformations correspond to

$$M_{i\ell_0\ell_1} = 2(b^{\ell_1} \delta_{i\ell_0} + b^{\ell_0} \delta_{i\ell_1} - b^i \delta_{\ell_0\ell_1}). \quad (109)$$

This is manifestly symmetric between ℓ_0 and ℓ_1 , and satisfies Eq. (100). We see that plugging this into Eq. (104) reproduces Eq. (81), and thus the special conformal consistency relation of Eq. 89 follows. Each of the 4 tensor symmetries come from an $M_{i\ell_0\ell_1}$ that is symmetric between ℓ_0 and ℓ_1 , fully traceless over any pair of indices, and transverse in the sense of Eq. (101). The corresponding $n = 1$ tensor consistency relation reads:

$$\lim_{\vec{q} \rightarrow 0} M_{i\ell_0\ell_1} \frac{\partial}{\partial q^{\ell_1}} \left[\frac{\langle \gamma_{i\ell_0}(\vec{q}) \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}}{P_\gamma(q)} \right] = -M_{i\ell_0\ell_1} \sum_{a=1}^N \left[\frac{1}{2} k_a^i \frac{\partial^2}{\partial k_a^{\ell_0} \partial k_a^{\ell_1}} \right] \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}, \quad (110)$$

where the dependence on $M_{i\ell_0\ell_1}$ can be removed by applying suitable projectors (see H2K).

For each $n \geq 2$, there are 4 purely tensor symmetries and 2 mixed symmetries where both π and γ

³¹Such a transformation would correspond to a time-independent rotation. See Appendix D though for the corresponding decaying mode, which although visible corresponds to a nonstandard choice of initial conditions.

transform nonlinearly. In general, any $n \geq 0$ consistency relation reads:

$$\begin{aligned}
& \lim_{\vec{q} \rightarrow 0} M_{i\ell_0\ell_1\dots\ell_n} \frac{\partial^n}{\partial q^{\ell_1}\dots\partial q^{\ell_n}} \left[\frac{\langle \gamma_{i\ell_0}(\vec{q}) \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}}{P_\gamma(q)} + \delta_{i\ell_0} \frac{D'(\eta)}{3c} \frac{\langle \pi_{\vec{q}} \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}}{P_\pi(q)} \right] \\
&= - M_{i\ell_0\ell_1\dots\ell_n} \sum_{a=1}^N \left[\delta_{i\ell_0} \frac{\partial^n}{\partial k_a^{\ell_1} \dots \partial k_a^{\ell_n}} - \delta_{i\ell_0} \frac{\delta_{n0}}{N} + \frac{k_a^i}{n+1} \frac{\partial^{n+1}}{\partial k_a^{\ell_0} \dots \partial k_a^{\ell_n}} \right. \\
&\quad + \delta_{i\ell_0} \frac{D(\eta_a)'}{3c} \left(\frac{\bar{\rho}'}{\bar{\rho}} \Big|_{\eta_a} + \partial_{\eta_a} \right) \frac{\partial^n}{\partial k_a^{\ell_1} \dots \partial k_a^{\ell_n}} \\
&\quad \left. - \frac{n}{c} \delta_{\ell_0\ell_1} (D(\eta_a) - D(\eta)(1 - \delta_{n1})) k_a^i \frac{\partial^{n-1}}{\partial k_a^{\ell_2} \dots \partial k_a^{\ell_n}} \right] \langle \delta_{\vec{k}_1} \dots \delta_{\vec{k}_N} \rangle^{c'}. \tag{111}
\end{aligned}$$

This is our most general result: each Newtonian gauge residual diffeomorphism described by Eq. (104) gives rise to a consistency relation given by Eq. (111).³² The consistency relations can also be written in a form in which the matrix M is projected out (see H2K). The familiar dilation and special conformal consistency relations are contained here and the M 's in those cases take a form that projects out the tensor term on the left hand side. Note that the soft modes are assumed to be at time η , while the hard modes are at time η_a for each momentum \vec{k}_a . Purely tensor consistency relations follow from those M 's that are fully traceless, hence setting the scalar contributions on the left hand side to be zero (and zeroing out terms proportional to the Kroenecker delta on the right hand side as well). For $n \geq 2$, there are choices of M (2 for each $n \geq 2$) that have a structure that gives both non-vanishing tensor and scalar contributions on the left hand side. It is worth pointing out that on the right hand side, the first set of terms (second line) are time-independent; they originate from the unitary gauge diffeomorphisms.³³ The second set of terms (third line) originate from ξ_{add}^0 , the additional time diffeomorphism that is necessary to keep us in Newtonian gauge. Likewise, the last set of terms (fourth line) come from ξ_{add}^i , and we have used the pure tensor consistency relation at level $(n-2)$ to move the terms proportional to $D(\eta)$ to the right hand side.

Let us study the taking of the Newtonian, i.e. sub-Hubble, limit. As explained in §3.3, it is helpful to rewrite the consistency relations using $\delta_{\vec{q}} \sim q^2 \pi_{\vec{q}} / \mathcal{H}$ (the precise relation is $\delta_{\vec{q}} = q^2 \pi_{\vec{q}} D / D'$ for the δ and π of pressureless matter in the sub-Hubble limit). Recalling that $c = D'' + 2\mathcal{H}D' \sim \mathcal{H}^2$, we see that Eq. (111) naïvely has a sub-Hubble limit of the schematic form:

$$\lim_{\vec{q} \rightarrow 0} \left\{ \frac{\mathcal{H}^2}{q^2} \frac{\langle \gamma_{\vec{q}} \delta_{\vec{k}} \dots \rangle^{c'}}{P_\gamma(q)} + \frac{\langle \delta_{\vec{q}} \delta_{\vec{k}_1} \dots \rangle^{c'}}{P_\delta(q)} \right\} \sim \left(\frac{q}{k} \right)^n \left(\frac{\mathcal{H}^2}{q^2} + n \frac{k^2}{q^2} \right) \langle \delta_{\vec{k}} \dots \rangle^{c'}, \tag{112}$$

³² With the exception of the dilation consistency relation ($n = 0$ with $M_{i\ell_0} \propto \delta_{i\ell_0}$), these consistency relations in general receive corrections on the right hand side which either involve replacing one of the hard modes by a hard (scalar or tensor) metric perturbation, or involve higher powers of the metric perturbations. These corrections arise because the associated diffeomorphisms generally need to be corrected order by order in metric perturbations (H2K, [16]). What we focus on in this paper are the lowest order terms in the diffeomorphisms (i.e. metric-independent contributions). Even in the nonlinear regime where density perturbations are large, the metric perturbations are in general small. Thus, the corrections to the consistency relations are negligible in applications where the hard modes are density (as opposed to metric) perturbations on sub-Hubble scales.

³³ The term $-\delta_{i\ell_0} \delta_{n0} / N$ arises from the removal of delta functions. See H2K for discussion.

where we have equated $\partial_q \sim 1/q$ and $\partial_k \sim 1/k$. Of the terms on the right hand side, the term suppressed in the sub-Hubble limit by (\mathcal{H}^2/q^2) arises from the second and third lines of Eq. (111), and the unsuppressed term comes from the last line of Eq. (111). It is also worth noting that the unsuppressed term is in general non-vanishing even if all the hard modes are at the same time, as long as the soft mode is at a different time; the $n = 1$ case (that gives rise to the KRPP consistency relation) is an exception rather than the rule.

At first sight, this suggests there is a non-trivial Newtonian limit for each $n > 0$, with the $n = 1$ case (KRPP) being one example. This is not the case because of the presence of tensor modes. In all $n \geq 2$ cases where the diffeomorphism generates a soft scalar, the same diffeomorphism generates a soft tensor as well. The tensor equation of motion $\gamma''_{ij} + 2\mathcal{H}\gamma'_{ij} - \nabla^2\gamma_{ij} = 0$ (Eq. 151) tells us that (1) ignoring ∇^2 , $\gamma_{ij} = \text{const.}$ is the growing mode solution (or more properly, the dominant mode solution; the other mode decays); (2) allowing for a small ∇^2 , the growing mode tensor solution gets corrected by a term proportional to D (see Appendix B); (3) when ∇^2 is important, the tensor mode oscillates with an amplitude that decays as $1/a$. Cases (1) and (2) pertain to super-Hubble modes while case (3) has to do with sub-Hubble ones. The purely tensor consistency relations follow from diffeomorphisms that are time-independent which generate tensor modes of type (1). The mixed scalar-tensor consistency relations follow from diffeomorphisms that generate tensor modes of type (2) (see footnote 39). In neither case are we allowed to take the soft tensor mode to within the Hubble radius. This is in contrast with the purely scalar consistency relations (such as dilation and special conformal transformation), where the time-dependence of the soft $\pi_{\vec{q}}$ remains the same whether it is outside or inside the Hubble radius. One might be tempted to say: within the Hubble radius, the tensor mode decays anyway, so why not just drop the tensor term from the consistency relations? This is not allowed because the tensor mode enters in both the numerator and denominator of $\langle \gamma_{\vec{q}...} \rangle' / P_\gamma(q)$. In general this ratio is independent of the amplitude of the tensor mode; the consistency relations express precisely this fact.

4 Discussion

Let us give a summary of our main results.

1. Consistency relations, between a squeezed $(N+1)$ -point function and an N -point function, are a generic consequence of nonlinearly realized symmetries, which shift the fields of interest by amounts that do not depend on the fields. They can be thought of as the LSS analogs of soft-pion theorems in particle physics, with the squeezed (soft) mode playing the role of a soft pion. The master formula for deriving consistency relations for any such nonlinearly realized symmetries is given in Eq. (22).³⁴

2. In this paper, we focus on nonlinearly realized symmetries that involve a change of coordinates; they are diffeomorphisms from the point of view of general relativity. These are residual diffeomorphisms that are allowed even after the usual gauge-fixing; they all share the property that they do not vanish at spatial infinity. From earlier work in the unitary gauge (or ζ -gauge, where the equal time surface is chosen to be one where the matter perturbation vanishes, and where for us ‘matter’ means the dark or pressureless matter), it is known that there is an infinite number of such symmetries, of the schematic form $x^i \rightarrow x^i + M^i x^{n+1}$, where $n = 0, 1, 2, \dots$ and the detailed index structures of M^i and x^{n+1} are

³⁴That is for the case of having a single pion; see Eq. (107) for an example that involves two independent soft pions (the scalar mode and the tensor mode(s)).

suppressed (H2K). Associated with these change of coordinates are both linear and nonlinear shifts in the fields, from which we derive consistency relations. For each n , the consistency relation (in unitary gauge) takes the schematic form:

$$\lim_{\vec{q} \rightarrow 0} \partial_q^n \left[\frac{\langle \zeta_{\vec{q}} \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^{c'}}{P_\zeta(q)} + \Theta(n > 1) \frac{\langle \gamma_{\vec{q}} \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^{c'}}{P_\gamma(q)} \right] \sim \partial_k^n \langle \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^{c'}, \quad (113)$$

where ζ and γ are the (soft) curvature and tensor perturbations, P_ζ and P_γ are their respective power spectra, and \mathcal{O} represents some observables of interest at hard momenta $\vec{k}_1, \dots, \vec{k}_N$. The correlation functions are connected and have the overall momentum-conserving delta functions removed. For $n \leq 1$, the tensor term on the left hand side is absent, hence the step function $\Theta(n > 1)$.

For LSS applications, especially when one is interested in sub-Hubble scales, it is useful to have the analogous relations in Newtonian gauge where the matter fluctuations are non-vanishing. We give the prescription for mapping each unitary gauge symmetry to its Newtonian gauge counterpart. This is given in Eq. (104). The diffeomorphism becomes $x^0 \rightarrow x^0 + \mathcal{H}^{-1} M x^n$ and $x^i \rightarrow x^i + M^i x^{n+1} + \mathcal{H}^{-2} M^i x^{n-1}$, where some of the suppressed indices of M need to be internally contracted; the factors of \mathcal{H}^{-1} and \mathcal{H}^{-2} are meant to indicate coefficients that are time-dependent, and the powers of Hubble reflect the order of magnitude of these coefficients. From these diffeomorphisms, the Newtonian gauge consistency relations read schematically:

$$\lim_{\vec{q} \rightarrow 0} \partial_q^n \left[\mathcal{H}^{-1} \frac{\langle \pi_{\vec{q}} \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^{c'}}{P_\pi(q)} + \Theta(n > 1) \frac{\langle \gamma_{\vec{q}} \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^{c'}}{P_\gamma(q)} \right] \sim [\partial_k^n + n \mathcal{H}^{-2} \partial_k^{n-2}] \langle \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^{c'}, \quad (114)$$

where $\pi_{\vec{q}}$ is the soft velocity potential. The precise form is given in Eq. (111). Because $\pi_{\vec{q}}$ is dimensionful (velocity $v^i = \partial_i \pi$), it is useful to rewrite this using the dimensionless $\tilde{\delta}_{\vec{q}} \equiv q^2 \pi_{\vec{q}} D / D' \sim q^2 \pi_{\vec{q}} / \mathcal{H}$.³⁵

$$\lim_{\vec{q} \rightarrow 0} \partial_q^n \left[q^2 \mathcal{H}^{-2} \frac{\langle \tilde{\delta}_{\vec{q}} \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^{c'}}{P_{\tilde{\delta}}(q)} + \Theta(n > 1) \frac{\langle \gamma_{\vec{q}} \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^{c'}}{P_\gamma(q)} \right] \sim [\partial_k^n + n \mathcal{H}^{-2} \partial_k^{n-2}] \langle \mathcal{O}_{\vec{k}_1} \dots \mathcal{O}_{\vec{k}_N} \rangle^{c'}. \quad (115)$$

Two noteworthy points: first, the right hand side should always be understood to contain corrections that vanish in the $\vec{q} \rightarrow 0$ limit i.e. we expect $O(q)$ corrections to the right; second, it appears that a non-trivial sub-Hubble limit exists (by sending $\mathcal{H} \rightarrow \infty$) for all $n > 0$, but there is a subtlety.

3. The subtlety has to do with the adiabatic mode condition. The consistency relations make three assumptions: the existence of nonlinearly realized symmetries, the single field initial condition, and the adiabatic mode condition. The last says that the soft mode generated nonlinearly by our symmetry transformation must satisfy the equation of motion at a low but finite momentum, i.e. the symmetry generated soft mode must have the correct time dependence to match that of a long wavelength physical mode. A corollary is that since each of our symmetries is a diffeomorphism, the *same* diffeomorphism must generate all the soft modes in the problem. This means that in a universe with multiple particle species, all the particles must move with the same velocity on large scales, implying that the equivalence principle is obeyed. (It is worth stressing that the equivalence principle needs to hold only on large

³⁵ We are careful to state this as a definition for $\tilde{\delta}_{\vec{q}}$ – it is equal to what we normally call $\delta_{\vec{q}}$ the matter density fluctuation only if \vec{q} is sub-Hubble.

scales; baryons or galaxies can move differently from dark matter on small scales.) This also means the sub-Hubble or Newtonian limit must be taken with care. Recall that the general relativistic consistency relations are strictly speaking $\vec{q} \rightarrow 0$ statements, and thus valid for super-Hubble scales $q < \mathcal{H}$. For the relations to continue to hold even as one takes $q > \mathcal{H}$ (keeping $k \gg q$ of course), the soft mode must maintain the same time dependence across the Hubble radius. The soft $\pi_{\vec{q}}$ has this property, and can be safely taken to be sub-Hubble. The same is not true for the soft tensor mode $\gamma_{\vec{q}}$. Thus, consistency relations which involve the tensor mode strictly hold only if the soft mode is super-Hubble. It is only in the special cases of $n = 0, 1$ (dilation and special conformal transformation) that the tensor term is absent from the left hand side, and the corresponding purely scalar consistency relations hold even within the Hubble radius (i.e. for $\mathcal{H} \ll q \ll k$). Of these two cases, $n = 0$ does not give a non-trivial right hand side; only $n = 1$ does in the sub-Hubble limit – this gives precisely the Newtonian consistency relation obtained by KRPP (Eq. 30). In fact, it is worth emphasizing that the KRPP relation as originally expressed (Eq. 33) should be thought of as containing two separate non-relativistic consistency relations: the lack of a $1/q^2$ pole is the sub-Hubble limit of the dilation relation and the leading $1/q$ term comes from the special conformal transformation, which (suitably generalized to Newtonian gauge) reduces to a time-dependent spatial translation in this limit.

4. The fact that the symmetries of interest are diffeomorphisms, albeit ones that do not vanish at spatial infinity, suggest the consistency relations are fairly robust: no detailed dynamical assumptions need to be made about the hard modes, and the only information we need is how they transform. They could be highly non-perturbative and even astrophysically complex, such as galaxy observables on small scales (though some care is needed if the hard modes are metric as opposed to density fluctuations, see footnote 32). We demonstrate this in the context of the Newtonian consistency relation by writing down explicitly a model of galaxy dynamics that allows for galaxy formation, mergers, dynamical friction and so on (§2.2). From the symmetry standpoint, the key assumption is that the dynamics and formation process of galaxies be frame invariant, i.e. aside from Hubble friction, the only possible dependence on velocity arises from gradients thereof or on velocity difference between different species. From the standpoint of the adiabatic mode condition, the key assumption is that all objects fall in the same way on large scales. From the practical standpoint, in terms of checking the consistency relation at a small momentum q (as opposed to a literally vanishing q), it is important to ensure q is sufficiently squeezed such that the soft modes evolve in a way consistent with the adiabatic mode condition. Checking the consistency relation observationally will allow us to test some fundamental assumptions in cosmology – in particular the single-field initial condition and the equivalence principle [18, 21].

5. We carry out a perturbative check on the robustness of the Newtonian consistency relation, by including a galaxy bias at the quadratic level (§2.2). For the consistency relation to hold, we find that any non-local quadratic bias cannot be too infrared divergent: supposing the galaxy density δ_g is related to the mass density δ by $\delta_{g\vec{k}} \sim b[\delta_{\vec{k}} + \int d^3k' (2\pi)^{-3} W(\vec{k}', \vec{k} - \vec{k}')(\delta_{\vec{k}'}\delta_{\vec{k}-\vec{k}'} - \langle \delta_{\vec{k}'}\delta_{\vec{k}-\vec{k}'} \rangle)]$, the quadratic kernel $W(\vec{k}', \vec{k} - \vec{k}')$ must grow slower than $1/k'$ in the small k' limit (Eq. 46). Thus, a local quadratic bias where W is independent of momenta, which is a form of bias commonly invoked, respects the consistency relation. Interestingly, the one case we know of where a $W \sim 1/k'$ behavior occurs is one where galaxies are born with a velocity bias – which is consistent with our understanding that having a large scale velocity bias violates the adiabatic mode condition [17]. The consistency relation allows the hard modes to be astrophysically complex observables such as the galaxy density δ_g . But the soft mode strictly

speaking should still be that of the dark (pressureless) matter. On the other hand, from an observational standpoint, galaxy density is much easier to measure. We show that the soft mode can be a galaxy observable provided that (1) the consistency relation is corrected by a multiplicative linear bias factor; (2) the galaxy power spectrum on large scales has a spectral slope of $n < 1$ (where $n \equiv d \ln P(q)/d \ln q$; see §2.2).

There are a number of outstanding questions for future investigations. The symmetries we are using are gauge symmetries: why is it that we manage to derive physically relevant statements out of gauge redundancies? We know consistency relations are not empty statements, because there are models that violate them, for instance models of inflation that involve multiple fields [4]. The answer presumably has to do with two aspects of the consistency relations: (1) they make certain physical assumptions about the initial condition, in particular single field initial condition; (2) while a strictly zero-momentum mode is unobservable, a soft mode with a small but non-zero momentum is observable, and consistency relations are ultimately statements about the dominant terms in a correlation function with a soft mode. It would be useful to clarify the curious role of gauge symmetries in physical statements. Further clarity in the derivation of the consistency relations is desirable for another reason: we see from our perturbative check in §2.2 that, even without galaxy bias, the validity of the consistency relations requires the soft power spectrum to be not too blue ($n < 3$). Why this should be so is not easy to see from the background wave argument. Can one see this from other arguments in the literature, such as the operator formalism (H2K), operator product expansion [8], or the effective action approach [11]?

Lastly, in this paper, we focus exclusively on nonlinearly realized symmetries that originate from diffeomorphisms. Could the LSS dynamics have other nonlinearly realized symmetries? Recently, it was pointed out by [22] and [23] that, indeed, further nonlinearly realized symmetries exist, albeit ones that involve the transformation of parameters as well. Can there be more?

Acknowledgements

We have benefited a great deal from discussions with Lasha Berezhiani, Paolo Creminelli, Justin Khoury, Alberto Nicolis, Jorge Noreña, Marko Simonović, and especially Filippo Vernizzi who collaborated with us on parts of the project. Our paper has a substantial overlap with the paper [16] by Paolo Creminelli, Jorge Noreña, Marko Simonović & Filippo Vernizzi. Our viewpoint is largely shaped by joint work with Justin Khoury & Kurt Hinterbichler, and Walter Goldberger & Alberto Nicolis. We further thank Guido D’Amico, Kurt Hinterbichler, Marco Peloso, Massimo Pietroni, Rafael Porto, Roman Scoccimarro, David Spergel and Matias Zaldarriaga for useful discussions. We thank David Spergel for organizing a non-Gaussianity workshop at Princeton University where there were many interesting discussions. We thank the Kavli Institute for Theoretical Physics, as well as the Lorentz Center at Universiteit Leiden, for hospitality while this work was in progress. This work is supported in part by the United States Department of Energy under DOE grant DE-FG02-92-ER40699, and by NASA under NASA ATP grant NNX10AN14G.

A A Lagrangian for Fluid with Pressure

In this Appendix, we connect the LSS Lagrangian discussed in §2.3 with a more commonly used fluid Lagrangian. We will continue to work within the zero vorticity regime. For generalizations to include a non-vanishing vorticity, see [30]. Let us consider the following action:

$$S = \int d^4x \sqrt{-g} \left[\frac{1}{2} M_P^2 R + \mathcal{P}(X) + \dots \right], \quad (116)$$

where the first term is the Einstein-Hilbert action, and $\mathcal{P}(X)$ is the fluid action, where \mathcal{P} is some function of $X \equiv \sqrt{-g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi}$ with ϕ describing the single degree of freedom of an irrotational fluid.³⁶ The ... stands for other possible matter or energy content in the universe, i.e. the background expansion need not be determined solely by the $\mathcal{P}(X)$ fluid in question. This is a completely relativistic action, and has been used by many authors [14, 15]. Our goal here is to take the non-relativistic limit, and connect the result with the action in §2.3 (Eq. 51).

We assume a metric of the form:

$$ds^2 = a^2 \left[-(1 + 2\Phi) d\eta^2 + (1 - 2\Psi) d\vec{x}^2 \right]. \quad (117)$$

The fluid energy-momentum $T_{\mu\nu}$ can be obtained from the fluid action by $\sqrt{-g} T_{\mu\nu} = -2\delta S_{\text{fluid}}/\delta g^{\mu\nu}$:

$$T_{\mu\nu} = 2\mathcal{P}_{,X} \partial_\mu \phi \partial_\nu \phi + g_{\mu\nu} \mathcal{P} \quad (118)$$

which is the energy-momentum of a perfect fluid, with the 4-velocity U^μ , energy density ρ and pressure P given by:

$$U^\mu = \frac{-\partial^\mu \phi}{\sqrt{X}}, \quad \rho = 2X\mathcal{P}_{,X} - \mathcal{P}, \quad P = \mathcal{P}. \quad (119)$$

A fluid with an equation of state $P = w\rho$ can be modeled by a $\mathcal{P}(X)$ of the form:

$$\mathcal{P}(X) \propto X^{\frac{1+w}{2w}}. \quad (120)$$

We are interested in the case of a small w . Let us split ϕ into a background $\bar{\phi}(\eta)$ and perturbation:

$$\phi = \bar{\phi} + \delta\phi, \quad \pi \equiv -\delta\phi/\bar{\phi}', \quad (121)$$

where we have defined π in terms of the field fluctuation $\delta\phi$. This definition is consistent with the interpretation of π as the velocity potential, as can be seen by working out U^μ in terms of ϕ and equating $U^\mu = a^{-1}(1, \vec{v})$ i.e. $v_i = \nabla_i \pi$ to the lowest order in perturbations. The background $\bar{\phi}$ obeys:

$$\partial_\eta(a^4 \bar{\rho}/\bar{\phi}') = 0, \quad (122)$$

which implies $\bar{\phi}' \propto a^{1-3w}$, using the fact that $\bar{\rho} \propto a^{-3(1+w)}$. We denote by \bar{X} the value of X evaluated at $\phi = \bar{\phi}$. Using the fact that $P = w\rho$, we find $1 + \delta = (1 + [\delta X/\bar{X}])^{(1+w)/2w}$ which implies

$$\frac{2w}{1+w} \ln(1 + \delta) = \ln\left(1 + \frac{\delta X}{\bar{X}}\right) \quad (123)$$

³⁶ The vanishing of vorticity can be expressed covariantly as $\epsilon^{\mu\nu\rho\sigma} u_\nu \partial_\rho u_\sigma = 0$.

Assuming both w and $\delta X/\bar{X}$ are small, but without assuming δ is small, we can approximate this by

$$2w \ln(1 + \delta) \sim \frac{\delta X}{\bar{X}}. \quad (124)$$

Let us write out $\delta X/\bar{X}$ explicitly in terms of the metric and ϕ fluctuations:

$$\frac{\delta X}{\bar{X}} = -2 \left[\Phi + \frac{1}{a}(a\pi)' + \frac{1}{2}(\nabla\pi)^2 \right], \quad (125)$$

where we have approximated $\bar{\phi}' \propto a$ (for small w), assumed $\Phi \sim v^2 \lesssim \pi' \sim \mathcal{H}\pi \ll 1$, and ignored terms to higher order (we regard Φ^2 , Φ' and $w\Phi$ as both higher order). Eqs. (124) and (125) combined give:

$$-w \ln(1 + \delta) = \Phi + \frac{1}{a}(a\pi)' + \frac{1}{2}(\nabla\pi)^2, \quad (126)$$

Applying the spatial gradient on this equation reproduces the Euler equation in the presence of pressure (Eq. 41), upon identifying w with c_s^2 , the sound speed squared.

We are interested in rewriting the action in Eq. (116) in terms of the fluctuations. In other words, we are not so much interested in the background as in the dynamics of the fluctuations. Thus, we ignore the background term in $\sqrt{-g}[M_P^2 R/2 + \mathcal{P}(X)]$. We also remove (tadpole) terms that are linear in fluctuations – they only serve to multiply the background equation of motion. Thus, we have (in the sub-Hubble, non-relativistic limit):

$$\begin{aligned} S &= S_{\text{EH}} + S_{\text{fluid}} \\ S_{\text{EH}} &= \int d^4x a^2 M_P^2 (\nabla_i \Psi \nabla_i \Psi - 2 \nabla_i \Psi \nabla_i \Phi) \\ S_{\text{fluid}} &= \int d^4x w a^4 \bar{\rho} \left[\frac{1+w}{2w} \frac{\delta X}{\bar{X}} + \left(\delta - \frac{1+w}{2w} \frac{\delta X}{\bar{X}} \right) \right], \end{aligned} \quad (127)$$

where S_{EH} comes from the Einstein-Hilbert action, and S_{fluid} comes from the fluid part of the action. For the latter, we have used the fact that $\mathcal{P}(X) = w\rho = w\bar{\rho}(1 + \delta)$, and removed the background piece $w\bar{\rho}$. We add and subtract $(1+w)\delta X/(2w\bar{X})$ to facilitate the removal of tadpole terms from expanding out $(1 + \delta) = (1 + \delta X/\bar{X})^{1+w/2w}$. We are to understand the last line as follows: the first $(1+w)\delta X/(2w\bar{X})$ should be understood to have the linear fluctuations removed, while the second $(1+w)\delta X/(2w\bar{X})$ has all terms in it.³⁷ The fluid part of S is therefore

$$S_{\text{fluid}} = \int d^4x a^4 \bar{\rho} \left[-\frac{1}{2}(\nabla\pi)^2 + F \right], \quad (128)$$

where F is

$$F \equiv w \left(\delta - \frac{1}{2w} \frac{\delta X}{\bar{X}} \right), \quad (129)$$

³⁷The determinant $\sqrt{-g}$ contains terms of order Φ and Φ^2 . Terms of order Φ multiplying the background are removed as tadpoles. Surviving terms can be seen to multiply at least one factor of w or of $v^2 \sim \Phi$ (the latter with no compensating $1/w$), and so are small compared to what we keep (which are of order v^2 or $v^2(v^2/w)$).

with δ and δX understood to be expressible in terms of Φ and π using Eqs. (124) and (125). We have already verified that the Euler equation with pressure holds (Eq. 126) – from the point of view of the action S_{fluid} , this merely serves as a definition for δ . Let us verify we obtain the Poisson and continuity equations by varying the action. First, we see that Ψ can be integrated out by setting $\Psi = \Phi$. In other words, let us work with the action:

$$S = \int d^4x \left\{ -a^2 M_P^2 (\nabla \Phi)^2 + a^4 \bar{\rho} \left[-\frac{1}{2} (\nabla \pi)^2 + F \right] \right\}. \quad (130)$$

The variation ΔF when we vary Φ , using Eqs. (124) and (125), is

$$\Delta F = \frac{1}{2} \frac{\Delta \delta X}{\bar{X}} \delta = -\delta \Delta \Phi, \quad (131)$$

giving us

$$\Delta S = \int d^4x [2a^2 M_P^2 \nabla^2 \Phi - a^4 \bar{\rho} \delta] \Delta \Phi, \quad (132)$$

and therefore the Poisson equation. This assumes that the only fluctuations sourcing Φ is from the $\mathcal{P}(X)$ fluid, which of course can be relaxed. The π equation of motion, on the other hand, follows from

$$\Delta F = \frac{1}{2} \frac{\Delta \delta X}{\bar{X}} \delta = - \left[\frac{1}{a} (a \Delta \pi)' + \nabla_i \pi \nabla \Delta \pi \right] \delta, \quad (133)$$

which together with the variation $\Delta (-\frac{1}{2} (\nabla \pi)^2)$ gives us the continuity equation $\delta' + \nabla_i [(1 + \delta) \nabla_i \pi] = 0$.

The action in Eq. (128) is a bit hard to use, because F involves a fairly nonlinear function of the fields. There are two possible simplifications.

We are interested in the $w \rightarrow 0$ limit. Keeping δ finite, Eq. (126) tells us

$$\Phi = -\frac{1}{a} \left[(a\pi)' + \frac{1}{2} a (\nabla \pi)^2 \right] \quad (134)$$

which is just the pressureless Euler equation again. Sending $F \rightarrow 0$, and substituting the above into Eq. (130), we obtain:

$$S = - \int d^4x \left[\frac{1}{2} \bar{\rho} a^4 (\nabla \pi)^2 + M_P^2 \left(\nabla \left[(a\pi)' + \frac{1}{2} a (\nabla \pi)^2 \right] \right)^2 \right] \quad (135)$$

reproducing Eq. (51) that we wrote down in §2.3. This justifies the normalization and sign that was adopted there.

The other possible simplification is to expand out $(1 + \delta) = (1 + \delta X / \bar{X})^{1+w/2w}$ to second order in $\delta X / \bar{X}$. We have resisted doing so earlier, because doing so effectively assumes $\delta X / \bar{X}$ is parametrically smaller than w (which is itself small). This is equivalent to assuming small δ , something we might not want to impose. It is nonetheless instructive to see what results:

$$S = \int d^4x \left\{ -a^2 M_P^2 (\nabla \Phi)^2 + \frac{1}{2} a^4 \bar{\rho} \left[-(\nabla \pi)^2 + c_s^{-2} \left(\frac{1}{a} (a\pi)' + \Phi + \frac{1}{2} (\nabla \pi)^2 \right)^2 \right] \right\}, \quad (136)$$

where we have set $w = c_s^2$. In the context of this action, we treat δ as *defined* by:

$$-c_s^2\delta = \Phi + \frac{1}{a}(a\pi)' + \frac{1}{2}(\nabla\pi)^2. \quad (137)$$

Notice how this differs from Eq. (126) in replacing $\ln(1+\delta)$ by δ on the left hand side. The reason for this definition is so that the Φ equation of motion gives the Poisson equation as usual. The π equation of motion can be seen to give the continuity equation. In other words, the full set of equations in this system are:

$$\delta' + \nabla_i[(1+\delta)v_i] = 0 \quad , \quad v'_i + v_j\nabla_j v_i + \mathcal{H}v_i = -\nabla_i\Phi - c_s^2\nabla_i\delta \quad , \quad \nabla^2\Phi = 4\pi G\bar{\rho}a^2\delta. \quad (138)$$

This is the set of equations one expects for a fluid with pressure, except the pressure term in the Euler equation is slightly modified from the non-perturbative one displayed in Eq. (41). Aside from this modification, this system of equations has the correct nonlinear structure. In particular, on length scales above the Jeans scale i.e. $k < k_J$ where $k_J^2 \equiv a^2\bar{\rho}/(2M_P^2c_s^2)$, one can ignore the pressure term, and the system reduces exactly to the standard pressureless LSS equations (Eq. 9). A useful feature of the action in Eq. (136) is that it shows clearly π has a kinetic term of the correct sign.

To summarize, the fluid action Eq. (130) gives the exact nonlinear equations for the perturbations of a fluid with pressure in the Newtonian limit. It simplifies in the zero-pressure limit to the action in Eq. (51), which gives the exact nonlinear equations for a pressureless fluid. It can be approximated by the action in Eq. (136) which gives a linearized pressure term for the Euler equation, but otherwise retains the full nonlinear structure of the exact theory.

B Derivation of the General Relativistic Adiabatic Mode Conditions in Newtonian Gauge

In this Appendix, we derive the adiabatic mode conditions appropriate for the Newtonian gauge, and derive the additional diffeomorphism laid out in Eq. (61).

For the purpose of deriving consistency relations, it is important that the modes generated nonlinearly by the symmetries be the low momentum limit of actual physical modes, i.e. they must obey adiabatic mode conditions (see §2.2). For the low momentum modes (and for them only), it is sufficient to consider the linearized Einstein equations, and study the time-dependence they imply for the perturbations. The linearized Einstein equations in Newtonian gauge are:

$$-\frac{1}{2}\delta G^0_0 = -4\pi G\delta T^0_0 \rightarrow \nabla^2\Psi - 3\mathcal{H}(\Psi' + \mathcal{H}\Phi) = 4\pi Ga^2\sum(\rho - \bar{\rho}), \quad (139)$$

$$\frac{1}{2}\delta G^0_i = 4\pi G\delta T^0_i \rightarrow -\partial_i(\Psi' + \mathcal{H}\Phi) + \frac{1}{4}\nabla^2 S_i = 4\pi Ga^2\sum(\bar{\rho} + \bar{P})(v_i + S_i), \quad (140)$$

$$\frac{1}{6}\delta G^k_k = \frac{4\pi}{3}G\delta T^k_k \rightarrow \Psi'' + \mathcal{H}(\Phi' + 2\Psi') + (2\mathcal{H}' + \mathcal{H}^2)\Phi - \frac{1}{3}\nabla^2(\Psi - \Phi) = 4\pi Ga^2\sum(P - \bar{P}), \quad (141)$$

$$\begin{aligned} \delta G^i_j - \frac{1}{3}\delta^i_j\delta G^k_k &= 8\pi G\left(\delta T^i_j - \frac{1}{3}\delta^i_j\delta T^k_k\right) \rightarrow \\ \left(\partial_i\partial_j - \frac{\delta_{ij}}{3}\nabla^2\right)(\Psi - \Phi) - (\partial_0 + 2\mathcal{H})\partial_{(i}S_{j)} + (\partial_0^2 + 2\mathcal{H}\partial_0 - \nabla^2)\frac{\gamma_{ij}}{2} &= 8\pi G\left(\delta T^i_j - \frac{1}{3}\delta^i_j\delta T^k_k\right). \end{aligned} \quad (142)$$

We have allowed the possibility that there might be multiple fluid components present (for instance dark matter, baryons, radiation, etc.), hence the summation on the right hand side, though we suppress the label for each component. Also useful are the linearized conservation equations, assuming each fluid is individually conserved. The continuity equation for each fluid component is

$$\delta'_n + \partial_i v_i = 3\Psi', \quad (143)$$

where δ_n is related to the density fluctuation $\delta \equiv (\rho - \bar{\rho})/\bar{\rho}$ by $(1+w)\delta_n = \delta$, with $w = P/\rho$ being the equation of state parameter of the fluid component of interest. This definition of δ_n is motivated by the fact that $\bar{\rho} \propto a^{-3(1+w)}$, and so it is $\bar{\rho}^{[1/(1+w)]}$ that redshifts like a^{-3} , i.e. one can think of $n \equiv \rho^{[1/(1+w)]}$ as the “number” density, and of δ_n as its fractional (small) fluctuation (for instance, for $w = 1/3$, n would be the number density of photons). In deriving Eq. (143), it is useful to know $\mathcal{H}^2 - \mathcal{H}' = 4\pi G a^2 \sum (\bar{\rho} + \bar{P})$. Note also that, in an analogous manner to Eq. (67):

$$\Delta_{\text{nl}} \delta_n = 3\mathcal{H}\xi^0. \quad (144)$$

The relativistic Euler equation for each component is:

$$(v_i + S_i)' + (1 - 3w)\mathcal{H}(v_i + S_i) = -\partial_i \Phi - w\partial_i \delta_n, \quad (145)$$

where we have assumed the fourth Einstein equation has a vanishing source.³⁸ Decomposing this last equation into scalar, vector and tensor parts, we have

$$\left(\partial_i \partial_j - \frac{\delta_{ij}}{3} \nabla^2 \right) (\Psi - \Phi) = 0, \quad (146)$$

$$(\partial_0 + 2\mathcal{H})\partial_{(i} S_{j)} = 0, \quad (147)$$

$$(\partial_0^2 + 2\mathcal{H}\partial_0 - \nabla^2)\gamma_{ij} = 0. \quad (148)$$

Following Weinberg [25], we demand that the (nonlinear part of the) symmetry-generated perturbations, as described in §3.1), solve the Einstein equations in a non-trivial way, that is, in a way that works even if we deform those perturbations slightly away from the zero momentum $q = 0$ limit. (See Eq. 14 for the Newtonian analog of this statement.) For scalar fluctuations, we therefore insist:

$$\text{scalar adiabatic mode condition : } \Psi = \Phi \quad , \quad -(\Psi' + \mathcal{H}\Phi) = 4\pi G a^2 \sum (\bar{\rho} + \bar{P})\pi, \quad (149)$$

which comes from removing spatial gradients from Eq. (146) and the scalar part of Eq. (140). Similarly, the vector adiabatic mode conditions are (if vector modes are present):

$$\text{vector adiabatic mode condition : } (\partial_0 + 2\mathcal{H})S_i = 0 \quad , \quad (v_i^\perp + S_i)' + (1 - 3w)\mathcal{H}(v_i^\perp + S_i) = 0, \quad (150)$$

which comes from Eq. (147) and the transverse component of Eq. (145) (v_i^\perp is the transverse part of the velocity v_i , i.e. the vorticity component). Note how the vector modes have only a single solution, which

³⁸This does not strictly hold if for instance the fluid is made out of a collection of relativistic particles, but it is a reasonably good approximation in LSS. In essence, we assume $\Phi = \Psi$, and $(a^2 S_i)' = 0$.

decays. Since single field inflation cannot generate vector modes, we will not consider them further here. The tensor adiabatic mode condition is simply the tensor equation of motion (148):

$$\text{tensor adiabatic mode condition} : \gamma''_{ij} + 2\mathcal{H}\gamma'_{ij} - \nabla^2\gamma_{ij} = 0. \quad (151)$$

Note that we do not wish to simply set the gradient to zero, because we are interested in diffeomorphisms generating a γ_{ij} that is the soft limit of a finite momentum physical mode.

Applying the above conditions to the (nonlinear part of the) symmetry-generated perturbations (Eqs. 60 and 65), we obtain:

$$\xi^{0'} + 2\mathcal{H}\xi^0 + \frac{1}{3}\partial_i\xi^i = 0 \quad , \quad \partial_i\xi^{i'} = 0 \quad , \quad \gamma_{ij} = -\left(\partial_i\xi^j + \partial_j\xi^i - \frac{2}{3}\delta_{ij}\partial_k\xi^k\right). \quad (152)$$

The first equality enforces $\Phi = \Psi$. The second equality enforces the second part of the scalar adiabatic mode condition, with the understanding that in the soft limit, all fluid components share the same velocity perturbation π . The third equality equates the tensor mode with the traceless part of the spatial metric generated by the diffeomorphism – this holds only if a certain gauge condition is satisfied such that the scalar contribution to the spatial metric resides entirely in its trace (see below). As far as the adiabatic mode condition is concerned, the important point is that γ_{ij} defined this way satisfies the tensor equation of motion (151). These three expressions constitute the adiabatic mode conditions on residual diffeomorphisms in Newtonian gauge. For a diffeomorphism to respect the Newtonian gauge, it must satisfy

$$\partial_i\xi^0 = \partial_0\xi^i \quad , \quad \nabla^2\xi^i + \frac{1}{3}\partial_i(\partial_k\xi^k) = 0 \quad (153)$$

such that $\Delta_{\text{nl}}g_{0i} = 0$ and the traceless part of $\Delta_{\text{nl}}g_{ij}$ is transverse (see Eq. 60).

As discussed in §3.1, one way to organize the set of Newtonian-gauge diffeomorphisms that satisfy Eqs. (152) and (153) is to relate each such diffeomorphism to a corresponding known residual diffeomorphism in the unitary gauge $\xi_{\text{unit.}}$ (Eq. 63):

$$\xi^0 = \xi_{\text{add.}}^0 \quad , \quad \xi^i = \xi_{\text{unit.}}^i + \xi_{\text{add.}}^i, \quad (154)$$

where the time-independent $\xi_{\text{unit.}}^i$ is supplemented by a time- and space- diffeomorphism $\xi_{\text{add.}}^0, \xi_{\text{add.}}^i$. The time-independent unitary-gauge diffeomorphism $\xi_{\text{unit.}}^i$ satisfies Eq. (57). Comparing this with Eq. (153), we see that $\xi_{\text{add.}}^i$ itself must satisfy the same: $\nabla^2\xi_{\text{add.}}^i + \partial_i(\partial_k\xi_{\text{add.}}^k)/3 = 0$. For this reason, we might as well absorb any time-independent part of $\xi_{\text{add.}}^i$ into the definition of $\xi_{\text{unit.}}^i$. From the second condition in Eq. (152), we see that $\partial_i\xi_{\text{add.}}^i$ must be independent of time. Suppose it is equal to some function $f(x)$. One can express $\xi_{\text{add.}}^i$ as a gradient and a curl (plus possibly some function that depends only on time). The divergence of the gradient is what matches up with $f(x)$, i.e. the gradient part is time-independent, and so by definition, it should have been absorbed into $\xi_{\text{unit.}}$ already. Thus, we can set $f(x) = 0$ and we can assume $\partial_i\xi_{\text{add.}}^i = 0$ without loss of generality. The first condition of Eq. (152) thus tells us

$$\xi_{\text{add.}}^{0'} + 2\mathcal{H}\xi_{\text{add.}}^0 + \frac{1}{3}\partial_i\xi_{\text{unit.}}^i = 0 \quad (155)$$

Recall from Eq. (65) that $\Delta_{\text{nl.}}\pi = \xi^0 = \xi_{\text{add.}}^0$. From Appendix C, we see that π in the soft limit has the time dependence of D' where $D(\eta)$ is the linear growth factor satisfying:

$$D'' + 2\mathcal{H}D' - c = 0 \quad (156)$$

where c is a constant. Comparing this against Eq. (155) and keeping only the growing solution, we see that

$$\xi_{\text{add.}}^0 = -\frac{1}{3c} D' \partial_i \xi_{\text{unit.}}^i, \quad (157)$$

confirming the time-diffeomorphism of Eq. (61). We can then solve for $\xi_{\text{add.}}^i$ from the first expression of Eq. (153) which tells us $\partial_i \xi_{\text{add.}}^0 = \partial_0 \xi_{\text{add.}}^i$, i.e.

$$\xi_{\text{add.}}^i = -\frac{1}{3c} D \partial_i (\partial_k \xi_{\text{unit.}}^k) = \frac{1}{c} D \nabla^2 \xi_{\text{unit.}}^i, \quad (158)$$

where the second equality follows from Eq. (57). This confirms the space-diffeomorphism of Eq. (61). As a self-consistency check, one can see that Eq. (57) also implies that $\partial_i \xi_{\text{add.}}^i = 0$. Lastly, it can also be checked that the tensor mode created by this diffeomorphism (the third expression of Eq. 152) obeys the tensor equation of motion. To see this, it is useful to note that because $\xi_{\text{unit.}}^i$ satisfies $\nabla^2 \xi_{\text{unit.}}^i + \partial_i (\partial_k \xi_{\text{unit.}}^k) / 3 = 0$, we also know $\nabla^2 \partial_k \xi_{\text{unit.}}^k = 0$, $\nabla^2 \partial_i \xi_{\text{unit.}}^j = -\partial_i \partial_j \partial_k \xi_{\text{unit.}}^k / 3$, and $\nabla^2 \nabla^2 \partial_i \xi_{\text{unit.}}^j = 0$.³⁹ It is worth noting that for pure tensor symmetries, where $\partial_i \xi_{\text{unit.}}^i = 0$, both $\xi_{\text{add.}}^0$ and $\xi_{\text{add.}}^i$ vanish, and so the pure tensor symmetries coincide in the Newtonian gauge and unitary gauge, as they should.

C Derivation of the General Relativistic Velocity Equation

Our goal in this Appendix is to derive the following equation for the velocity potential π :

$$(\pi' + 2\mathcal{H}\pi - C)' - 3w\mathcal{H}(\pi' + 2\mathcal{H}\pi - C) = w(g' + \mathcal{H}g) - (1 + 3w) \left[(\mathcal{H}^2 - \mathcal{H}')\pi - 4\pi G a^2 \sum (\bar{\rho} + \bar{P})\pi \right], \quad (159)$$

where $g \equiv \int d\eta \nabla^2 \pi$, and C is a constant in time but not space (determined by initial conditions). Here, π refers to the velocity potential of some particular fluid component of interest with an equation of state parameter w – except in the very last term where $\sum (\bar{\rho} + \bar{P})\pi$ refers to a sum over all fluid components. We will use this equation to deduce useful statements about the time-dependence of π in the soft limit.

The continuity equation (143) can be integrated once to obtain:

$$\delta_n = 3(\Psi + C) - g, \quad g \equiv \int d\eta \nabla^2 \pi, \quad (160)$$

where C denotes some integration constant – independent of time, but dependent on space in general. This can be substituted into (the scalar part of) the relativistic Euler equation (145) and integrated once to give

$$\pi' + \mathcal{H}(1 - 3w)\pi = -(1 + 3w)\Psi - 3wC + wg. \quad (161)$$

³⁹ In other words, the combined action of $\xi_{\text{unit.}}^i + \xi_{\text{add.}}^i$ generates a tensor mode of the form $\gamma_{ij} = [1 + (D/c)\nabla^2]\gamma_{ij\text{unit.}}$ where $\gamma_{ij\text{unit.}}$ is the tensor mode generated by the time-independent unitary diffeomorphism alone. That the constant tensor (growing) mode gets corrected at finite momentum by a term proportional to momentum squared should not be surprising. The time dependence can also be checked explicitly by solving the tensor equation of motion in the small but finite momentum limit.

On the other hand, (the scalar part of) the δG^0_i equation (140) can be integrated once to obtain

$$-(\Psi' + \mathcal{H}\Psi) = 4\pi G a^2 \sum (\bar{\rho} + \bar{P})\pi, \quad (162)$$

where we have assumed $\Psi = \Phi$. One can solve for Ψ from Eq. (161), substitute the result into Eq. (162), and subtract from both sides $(\mathcal{H}^2 - \mathcal{H}')\pi$. This gives Eq. (159). Note that in this derivation, we have not thrown away any gradient terms, i.e. we have not made any super-Hubble approximation.

Equation (159) simplifies if π happens to be the same for all fluid components, in which case what appears within the square brackets $[]$ sums to zero, by virtue of $\mathcal{H}^2 - \mathcal{H}' = 4\pi G a^2 \sum (\bar{\rho} + \bar{P})$. This happens, for instance, if we work on super-Hubble scales and assume adiabatic initial conditions. One can check that this is a self-consistent solution on super-Hubble scales, and assuming all fluid components move with the same π , the entire right hand side of Eq. (159) vanishes, implying:

$$\pi' + 2\mathcal{H}\pi - C \propto a^{3w}. \quad (163)$$

This suggests different fluid components (with different w 's) evolve differently, *unless* the proportionality constant is in fact zero, i.e.

$$\pi' + 2\mathcal{H}\pi - C = 0. \quad (164)$$

With this choice of the initial condition, it is thus consistent to have the same π for all fluid components on super-Hubble scales. *Interestingly, for pressureless matter ($w = 0$), Eq. (164) holds even on sub-Hubble (but linear) scales, after radiation domination.* This can be seen by setting $w = 0$ in Eq. (159), and noting that during matter or cosmological constant domination, the terms within the square brackets $[]$ still sum to zero. This means that for a wave-mode (of pressureless matter) that enters the Hubble radius after radiation domination, Eq. (164) holds for its entire history. For a wave-mode that enters the Hubble radius before matter domination, however, Eq. (164) does not hold in the intermediate period when the mode is within the Hubble radius during the radiation dominated phase.⁴⁰

As we see in §3.2, the fact that Eq. (164) holds for pressureless matter both inside and outside the Hubble radius (as long as the wave-mode of interest crosses the Hubble radius after radiation domination) enables us to have interesting consistency relations in the Newtonian limit. It is also worth noting that since $\partial_i \pi$ describes the dark matter velocity on all (linear) scales, including sub-Hubble ones, where we know the velocity scales with time as D' (D being the linear growth factor), we expect

$$D'' + 2\mathcal{H}D' - c = 0, \quad (165)$$

where c is some constant whose normalization is arbitrary – its normalization is tied to the normalization of the growth factor D . That this relation holds for the Newtonian growth factor in a matter dominated universe is easy to check: $D \propto a$. That this is true for more general cases is less familiar. Let us check this for a universe with a cosmological constant.

For a flat universe with pressureless matter and a cosmological constant, the linear growth factor can be written in closed form [32]:

$$D = \frac{5\Omega_m^0}{2} \frac{H(a)}{H_0} \int_0^a \frac{d\tilde{a}}{(\tilde{a}H(\tilde{a})/H_0)^3}, \quad (166)$$

⁴⁰ It is worth pointing out that Eq. (164), when substituted into Eq. (161) gives $\Psi = -(\pi' + \mathcal{H}\pi)$ – this holds as long as the wg term can be ignored, which can be justified either for super-Hubble scales, or for $w = 0$.

where Ω_m^0 is the matter density today, H_0 is the Hubble constant today, and the normalization is chosen such that D equals the scale factor a in the early universe. Note that $H = a'/a^2$ whereas $\mathcal{H} = a'/a$.

Let us rewrite what we want to show, Eq. (165), as follows:

$$H \frac{d}{da} \left(a^4 H \frac{dD}{da} \right) = c. \quad (167)$$

First, note that

$$\frac{d}{da} \left(\frac{H}{H_0} \right) = -\frac{3\Omega_m^0}{2a^4} \frac{H_0}{H}. \quad (168)$$

We can therefore work out:

$$\frac{dD}{da} = -\frac{3\Omega_m^0}{2a^4} \frac{H_0}{H} \frac{5\Omega_m^0}{2} \int_0^a \frac{d\tilde{a}}{(\tilde{a}H(\tilde{a})/H_0)^3} + \frac{5\Omega_m^0}{2} \frac{H_0^2}{a^3 H^2}. \quad (169)$$

Therefore,

$$\frac{d}{da} \left(a^4 H \frac{dD}{da} \right) = \frac{5\Omega_m^0}{2} \frac{H_0^2}{H}, \quad (170)$$

implying the desired result Eq. (167).

D Decaying Modes and Vector Symmetries

At several places in our discussion we have omitted the second solution to the scalar and tensor equations of motion, which decays at late times. It is worth discussing the decaying solution in a little more detail, both for completeness and because the reasons for ignoring it are slightly subtle.

The decaying solution arises at the linearized level because both the scalar and tensor equations of motion for the linearized modes (Eq. 13 or Eq. 165, and 151) are second order. In particular, we note that Eq. 165 for the linear growth factor $D(\eta)$ can allow for a decaying piece $D' \propto 1/a^2$ which is independent of c as well as the strictly growing piece (which does depend on having $c \neq 0$, and which is what is usually meant by the linear growth factor). The decaying piece corresponds to varying the lower limit of the integral in Eq. 166, which is arbitrary. It is straightforward to check that this solution gives the correct decaying solution $\delta \propto \mathcal{H}/a$ for Eq. 13 in the Newtonian limit, using the helpful lemma

$$\mathcal{H}' - \mathcal{H}^2 = -4\pi G_N a^2 \Sigma(\bar{\rho} + \bar{P}) = -4\pi G_N \bar{\rho}_{\text{matter}} \propto \frac{1}{a^2} \quad (171)$$

in a Λ CDM universe. From Eq. 155, the most general diffeomorphism $\xi_{\text{add.}}^0, \xi_{\text{unit.}}^i + \xi_{\text{add.}}^i$ involving scalar and tensor modes allowed by the adiabatic mode conditions includes the decaying piece

$$\xi_{\text{add.,decay}}^0 = \frac{d(x)}{a^2}, \quad \xi_{\text{add.,decay}}^i = \int^\eta \frac{\partial^i d(x)}{a^2} \quad (172)$$

where $d(x)$ is harmonic because of Eqs. (152) and (153). It is straightforward to check that this satisfies the tensor equation 151 in the super-Hubble limit as well.

Note that the decaying mode is independent of the time-independent (growing mode) unitary transformation, and so the diffeomorphism corresponding to a decaying mode is a separate symmetry. We can Taylor expand as in §3.4 to write the most general decaying symmetries as

$$\xi_{\text{decay}}^0 = \frac{1}{n!} \frac{1}{a^2} M_{\ell\ell\ell_1\ldots\ell_n} x^{\ell_1} \cdots x^{\ell_n}, \quad \xi_{\text{decay}}^i = \frac{1}{(n-1)!} \int^\eta \frac{1}{a^2} M_{\ell\ell\ell_2\ldots\ell_n}^i x^{\ell_2} \cdots x^{\ell_n}. \quad (173)$$

for $(n \geq 1)$, where the M 's are constant and obey the usual transversality and adiabatic transversality conditions.

Can we derive consistency relations for the decaying modes using these symmetries, using Eq. 22 or its generalization Eq. 105? We argue that the answer is no, though it is not enough to say that these simply decay away. Rather, keeping the decaying modes would correspond to a nonstandard choice of the initial vacuum state in the far past: if the decaying mode is not set to zero the energy associated with these modes (the scalar part of the action scales like $\rho a^4 \sim a^2 (\nabla \Phi)^2 \sim \mathcal{H}^2 / a^2$) becomes divergent at early times. Had we chosen to ignore this problem and work within the putative vacuum containing only decaying modes, we could have, in which case the lack of the time-dependent piece would make our consistency relations look slightly different from Eq. 111. Note that it is not true that the consistency relations should vanish in this case because the modes decay at late times: this is because the consistency relation is a ratio between the $(N+1)$ -pt function and the power spectrum, both of which decay, but the ratio on the right hand side does not.

For the sake of completeness, we discuss the case where the symmetries may involve vector modes – unlike the scalars and tensors, these have only a decaying solution. The condition

$$\nabla^2 \xi^i + \frac{1}{3} \partial^i (\partial_j \xi^j) = 0. \quad (174)$$

on the spatial part of the diffeomorphism will still be obeyed, but the condition $\partial_0 \xi^i = \partial_i \xi^0$ will be violated and replaced by the weaker condition

$$\nabla^2 \xi^0 = \partial^0 \partial_i \xi^i = 0. \quad (175)$$

where in the second equality we have made use of the second equation in Eq. 152. Using the vector adiabatic mode conditions (Eq. 150) we have

$$(\partial_i \xi^0 - \partial_0 \xi^i)_{\text{vec.}} = \frac{\bar{\xi}^i}{a^2}, \quad (\partial_i \xi^0)_{\text{vec.}} \propto \frac{1}{a} \quad (176)$$

where $\bar{\xi}^i$ is transverse and time-independent. The second of these conditions is clearly incompatible with the first condition in Eq. 152:

$$\xi^{0'} + 2\mathcal{H}\xi^0 + \frac{1}{3}\partial_i \xi^i = 0, \quad (177)$$

unless $(\partial_i \xi^0)_{\text{vec.}}$ vanishes, and so the vector part of the symmetry will be $\bar{\xi}^i \int \frac{d\eta}{a^2} \subset \xi^i$. We can Taylor expand

$$\bar{\xi}^i = \sum \frac{1}{(n+1)!} \bar{M}_{\ell_0\ell_1\ldots\ell_n}^i x^{\ell_0} \cdots x^{\ell_n} \quad (178)$$

The \bar{M} 's are completely traceless, and they obey the usual tensor transversality conditions Eqs. 100, 101 as well, so these are vector-tensor symmetries. Note that they obey the tensor equation of motion Eq.

151 on superhorizon scales, though they correspond to the decaying mode solution. For $n \geq 1$ there will be 4 such symmetries at each level. For $n = 0$, there are additional symmetries where with $\bar{M}_{i\ell_0}$ is antisymmetric in the indices; these correspond to time-dependent rotations.⁴¹ They will obey the further adiabatic transversality condition

$$\hat{q}^i(\bar{M}_{i\ell_0}(\hat{q}) - \bar{M}_{\ell_0 i}) = 0 \quad (179)$$

which will reduce the number of allowed polarizations from 3 to 2. Since a localized rotation necessarily involves shearing, we need this condition to enforce transversality in addition to the antisymmetric tensor structure.

To summarize, at $n = 0$ there are two purely vector symmetries, and for $n \geq 1$ there are four vector + tensor symmetries. Since vector modes always decay, for our choice of vacuum there are no consistency relations that involve vector modes.

Bibliography

- [1] A. Kehagias and A. Riotto, “Symmetries and Consistency Relations in the Large Scale Structure of the Universe,” Nucl. Phys. B **873**, 514 (2013) [arXiv:1302.0130 [astro-ph.CO]].
- [2] M. Peloso and M. Pietroni, “Galilean invariance and the consistency relation for the nonlinear squeezed bispectrum of large scale structure,” JCAP **1305**, 031 (2013) [arXiv:1302.0223 [astro-ph.CO]].
- [3] J. M. Maldacena, “Non-Gaussian features of primordial fluctuations in single field inflationary models,” JHEP **0305**, 013 (2003) [astro-ph/0210603].
- [4] P. Creminelli and M. Zaldarriaga, “Single field consistency relation for the 3-point function,” JCAP **0410**, 006 (2004) [astro-ph/0407059].
- [5] P. Creminelli, J. Norena and M. Simonović, “Conformal consistency relations for single-field inflation,” JCAP **1207**, 052 (2012) [arXiv:1203.4595 [hep-th]].
- [6] K. Hinterbichler, L. Hui and J. Khoury, “Conformal Symmetries of Adiabatic Modes in Cosmology,” JCAP **1208**, 017 (2012) [arXiv:1203.6351 [hep-th]].
- [7] (H2K) K. Hinterbichler, L. Hui and J. Khoury, “An Infinite Set of Ward Identities for Adiabatic Modes in Cosmology,” arXiv:1304.5527 [hep-th].
- [8] V. Assassi, D. Baumann and D. Green, “On Soft Limits of Inflationary Correlation Functions,” JCAP **1211**, 047 (2012) [arXiv:1204.4207 [hep-th]].
- [9] V. Assassi, D. Baumann and D. Green, “Symmetries and Loops in Inflation,” JHEP **1302**, 151 (2013) [arXiv:1210.7792 [hep-th]].
- [10] A. Kehagias and A. Riotto, Nucl. Phys. B **864**, 492 (2012) [arXiv:1205.1523 [hep-th]].

⁴¹Note that for $n \geq 1$, the tensor structure that is antisymmetric in the first two indices and symmetric in the last $n + 1$ will vanish identically.

- [11] W. D. Goldberger, L. Hui and A. Nicolis, “One-particle-irreducible consistency relations for cosmological perturbations,” *Phys. Rev. D* **87**, 103520 (2013) [arXiv:1303.1193 [hep-th]].
- [12] G. L. Pimentel, “Inflationary Consistency Conditions from a Wavefunctional Perspective,” arXiv:1309.1793 [hep-th].
- [13] L. Berezhiani and J. Khoury, “Slavnov-Taylor Identities for Primordial Perturbations,” arXiv:1309.4461 [hep-th].
- [14] L. Boubekur, P. Creminelli, J. Norena and F. Vernizzi, “Action approach to cosmological perturbations: the 2nd order metric in matter dominance,” *JCAP* **0808**, 028 (2008) [arXiv:0806.1016 [astro-ph]].
- [15] P. Creminelli, G. D’Amico, J. Norena and F. Vernizzi, “The Effective Theory of Quintessence: the w_j-1 Side Unveiled,” *JCAP* **0902**, 018 (2009) [arXiv:0811.0827 [astro-ph]].
- [16] P. Creminelli, J. Norena, M. Simonović and F. Vernizzi, “Single-Field Consistency Relations of Large Scale Structure,” arXiv:1309.3557 [astro-ph.CO].
- [17] M. Peloso and M. Pietroni, “Ward identities and consistency relations for the large scale structure with multiple species,” arXiv:1310.7915 [astro-ph.CO].
- [18] A. Kehagias, J. Norena, H. Perrier and A. Riotto, “Consequences of Symmetries and Consistency Relations in the Large-Scale Structure of the Universe for Non-local bias and Modified Gravity,” arXiv:1311.0786 [astro-ph.CO].
- [19] A. Kehagias, H. Perrier and A. Riotto, “Equal-time Consistency Relations in the Large-Scale Structure of the Universe,” arXiv:1311.5524 [astro-ph.CO].
- [20] P. Creminelli, Jrm. Gleyzes, M. Simonović and F. Vernizzi, “Single-Field Consistency Relations of Large Scale Structure. Part II: Resummation and Redshift Space,” arXiv:1311.0290 [astro-ph.CO].
- [21] P. Creminelli, Jrm. Gleyzes, L. Hui, M. Simonović and F. Vernizzi, “Single-Field Consistency Relations of Large Scale Structure. Part III: Test of the Equivalence Principle,” arXiv:1312.6074 [astro-ph.CO].
- [22] P. Valageas, “Angular averaged consistency relations of large-scale structures,” arXiv:1311.4286 [astro-ph.CO].
- [23] A. Kehagias, H. Perrier and A. Riotto, “Equal-time Consistency Relations in the Large-Scale Structure of the Universe,” arXiv:1311.5524 [astro-ph.CO].
- [24] P. Valageas, “Consistency relations of large-scale structures,” arXiv:1311.1236 [astro-ph.CO].
- [25] S. Weinberg, “Adiabatic modes in cosmology,” *Phys. Rev. D* **67**, 123504 (2003) [astro-ph/0302326].
- [26] L. Hui, A. Nicolis and C. Stubbs, “Equivalence principle implications of modified gravity models,” *Phys. Rev. D* **80** 104002 (2009) [arXiv:0905.2966 [astro-ph.CO]].

- [27] L. Hui and A. Nicolis, “An Equivalence principle for scalar forces,” *Phys. Rev. Lett.* **105**, 231101 (2010) [arXiv:1009.2520 [hep-th]].
- [28] R. K. Sheth, K. C. Chan and R. Scoccimarro, “Non-local Lagrangian bias,” *Phys. Rev. D* **87**, 083002 (2013) [arXiv:1207.7117 [astro-ph.CO]].
- [29] K. C. Chan, R. Scoccimarro and R. K. Sheth, “Gravity and Large-Scale Non-local Bias,” *Phys. Rev. D* **85**, 083509 (2012) [arXiv:1201.3614 [astro-ph.CO]].
- [30] S. Dubovsky, T. Gregoire, A. Nicolis and R. Rattazzi, “Null energy condition and superluminal propagation,” *JHEP* **0603**, 025 (2006) [hep-th/0512260].
- [31] M. Crocce and R. Scoccimarro, “Renormalized cosmological perturbation theory,” *Phys. Rev. D* **73**, 063519 (2006) [astro-ph/0509418].
- [32] S. Dodelson, “Modern cosmology,” Amsterdam, Netherlands: Academic Pr. (2003), 440 pp.